

US008222482B2

(12) **United States Patent**  
**Bobzin et al.**

(10) **Patent No.:** **US 8,222,482 B2**  
(45) **Date of Patent:** **Jul. 17, 2012**

(54) **MODULATING PLANT OIL LEVELS**

(75) Inventors: **Steven Craig Bobzin**, Malibu, CA (US);  
**Daniel Mumenthaler**, Bonita, CA (US);  
**Boris Jankowski**, Newbury Park, CA  
(US); **Joel Cruz Rarang**, Granada Hills,  
CA (US)

(73) Assignee: **Ceres, Inc.**, Thousand Oaks, CA (US)

(\*) Notice: Subject to any disclaimer, the term of this  
patent is extended or adjusted under 35  
U.S.C. 154(b) by 656 days.

(21) Appl. No.: **12/161,935**

(22) PCT Filed: **Jan. 26, 2007**

(86) PCT No.: **PCT/US2007/002214**

§ 371 (c)(1),  
(2), (4) Date: **Apr. 15, 2009**

(87) PCT Pub. No.: **WO2007/089610**

PCT Pub. Date: **Aug. 9, 2007**

(65) **Prior Publication Data**

US 2009/0324797 A1 Dec. 31, 2009

**Related U.S. Application Data**

(60) Provisional application No. 60/762,422, filed on Jan.  
26, 2006, provisional application No. 60/797,077,  
filed on May 1, 2006.

(51) **Int. Cl.**

**A01H 1/00** (2006.01)  
**A01H 5/00** (2006.01)  
**A01H 5/10** (2006.01)  
**C12N 5/14** (2006.01)

(52) **U.S. Cl.** ..... **800/278; 800/281; 800/295; 800/298;**  
**800/306; 800/310; 800/312; 800/314; 800/322;**  
**800/320.1; 800/320.2; 435/410; 435/419**

(58) **Field of Classification Search** ..... None  
See application file for complete search history.

(56) **References Cited**

**U.S. PATENT DOCUMENTS**

4,987,071 A 1/1991 Cech et al.  
5,034,323 A 7/1991 Jorgensen et al.  
5,204,253 A 4/1993 Sanford et al.  
5,231,020 A 7/1993 Jorgensen et al.  
5,254,678 A 10/1993 Haseloff et al.  
5,538,880 A 7/1996 Lundquist et al.  
5,766,847 A 6/1998 Jackle et al.  
5,878,215 A 3/1999 Kling et al.  
5,945,306 A 8/1999 Bandman et al.  
5,998,700 A 12/1999 Lightfoot et al.  
6,004,753 A 12/1999 Yue et al.  
6,013,863 A 1/2000 Lundquist et al.  
6,287,778 B1 9/2001 Huang et al.  
6,319,713 B1 11/2001 Patten et al.  
6,326,527 B1 12/2001 Kirihara et al.  
6,329,571 B1 12/2001 Hiei  
6,417,429 B1 7/2002 Hein et al.  
6,423,885 B1 7/2002 Waterhouse et al.

6,452,067 B1 9/2002 Bedbrook et al.  
6,455,253 B1 9/2002 Patten et al.  
6,500,614 B1 12/2002 Arguello et al.  
6,521,453 B1 2/2003 Crameri et al.  
6,573,099 B2 6/2003 Graham  
6,648,930 B2\* 11/2003 Ulrich et al. .... 44/308  
6,753,139 B1 6/2004 Baulcombe et al.  
6,777,588 B2 8/2004 Waterhouse et al.  
6,906,244 B2 6/2005 Fischer et al.  
7,173,121 B2 2/2007 Fang  
7,214,789 B2 5/2007 Pennell  
7,312,376 B2 12/2007 Apuya et al.  
7,378,571 B2 5/2008 Apuya  
7,402,667 B2 7/2008 Cook et al.  
7,429,692 B2 9/2008 Dang  
7,598,367 B2 10/2009 Cook et al.  
2003/0175783 A1 9/2003 Waterhouse et al.  
2003/0175965 A1 9/2003 Lowe et al.  
2003/0180945 A1 9/2003 Wang et al.  
2003/0233670 A1 12/2003 Edgerton et al.  
2004/0137466 A1 7/2004 Jofuku et al.  
2004/0214330 A1 10/2004 Waterhouse et al.  
2005/0108791 A1\* 5/2005 Edgerton ..... 800/284  
2006/0021083 A1 1/2006 Cook  
2006/0041952 A1 2/2006 Cook  
2006/0260004 A1 11/2006 Fang et al.  
2007/0006335 A1 1/2007 Cook

**FOREIGN PATENT DOCUMENTS**

EP 0 511 979 8/1994  
EP 292 435 7/1995  
EP 0 513 849 8/1998  
EP 0 465 875 11/1998  
EP 0 534 858 4/2000  
JP 2004 121047 4/2004  
WO 95/06128 3/1995  
WO 97/01952 1/1997

(Continued)

**OTHER PUBLICATIONS**

U.S. Appl. No. 60/505,689, filed Sep. 23, 2003, Cook.  
U.S. Appl. No. 60/518,075, filed Nov. 6, 2003, Pennell.  
U.S. Appl. No. 60/544,771, filed Feb. 13, 2004, Cook.  
U.S. Appl. No. 60/558,869, filed Apr. 1, 2004, Cook.  
U.S. Appl. No. 60/583,609, filed Jun. 30, 2004, Alexandrov.  
U.S. Appl. No. 60/583,691, filed Jun. 30, 2004, Alexandrov.  
U.S. Appl. No. 60/612,891, filed Sep. 23, 2004, Kwok.  
U.S. Appl. No. 60/637,140, filed Dec. 16, 2004, Feldman.  
U.S. Appl. No. 60/757,544, filed Jan. 9, 2006, Dang.  
U.S. Appl. No. 60/776,307, filed Feb. 24, 2006, Kwok.  
Abler et al. "Isolation and characterization of a genomic sequence  
encoding the maize Cat3 catalase gene" *Plant Mol. Biol.*, 22:1031-  
1038 (1993).

(Continued)

*Primary Examiner* — Eileen B O Hara

(74) *Attorney, Agent, or Firm* — Fish & Richardson P.C.

(57) **ABSTRACT**

Methods and materials for modulating (e.g., increasing or  
decreasing) oil levels in plants are disclosed. For example,  
nucleic acids encoding oil-modulating polypeptides are dis-  
closed as well as methods for using such nucleic acids to  
transform plant cells. Also disclosed are plants having  
increased oil levels and plant products produced from plants  
having increased oil levels.

**55 Claims, 25 Drawing Sheets**

## FOREIGN PATENT DOCUMENTS

WO	98/36083	8/1998
WO	98/53083	11/1998
WO	99/32619	7/1999
WO	99/53030	10/1999
WO	00/31281	6/2000
WO	01/18191	3/2001
WO	01/35725	5/2001
WO	01/75164	11/2001
WO	02/10210	2/2002
WO	02/15675	2/2002
WO	02/16655	2/2002
WO	02/46449	6/2002
WO	02/081714	10/2002
WO	03/013227	2/2003
WO	03/095654	11/2003
WO	2005/011105	2/2005
WO	2005/023639	3/2005
WO	2005/034308	4/2005
WO	2005/098007	10/2005
WO	2006/005023	1/2006
WO	2006/034479	3/2006

## OTHER PUBLICATIONS

- Ahn et al., "Homoeologous relationships of rice, wheat and maize chromosomes" *Molecular and General Genetics*, 241:483-490 (1993).
- Alonso-Blanco et al., "The use of recombinant inbred lines (RILs) for genetic mapping," In *Methods in Molecular Biology* (J.M. Martinez-Zapater and J. Salinas, Humana Press, Totowa, NJ., 1998), 82: 137-146.
- An et al., "New cloning vehicles for transformation of higher plants," *The EMBO Journal*, 1985, 4(2): 277-284.
- Apuya et al., "RASPBERRY3 Gene encodes a Novel Protein Important for Embryo Development," *Plant Physiology*, 2002, 129(2):691-705.
- Armaleo et al., "Biolistic nuclear transformation of *Saccharomyces cerevisiae* and other fungi," *Current Genetics*, 1990, 17: 97-103.
- Baerson et al., "Developmental regulation of an acyl carrier protein gene promoter in vegetative and reproductive tissues" *Plant Mol. Biol.*, 22(2):255-267 (1993).
- Bateman et al., "Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins" *Nucl. Acids Res.*, 27(1):260-262 (1999).
- Baud et al., "An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS," *Plant Physiol. Biochem.*, 2002, 40:151-160.
- Bechtold et al., "In planta *Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants" *C.R. Acad. Sci. Paris*, 316:1194-1199 (1993).
- Bonner et al., "Reduction in the rate of DNA reassociation by sequence divergence" *J. Mol. Biol.*, 81:123-135 (1973).
- Bradshaw et al., "A new vector for recombination-based cloning of large DNA fragments from yeast artificial chromosomes," *Nucl. Acids. Rec.*, 1995, 23(23): 4850-4856.
- Broothaerts et al., "Gene transfer to plants by diverse species of bacteria," *Nature*, 2005, 433: 629-633.
- Brummell, et al., "Inverted repeat of a heterologous 3'-untranslated region for high-efficiency, high-throughput gene silencing" *Plant J.* 33:793-800 (2003).
- Burke et al., "Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors," *Science*, 1987, 236: 806-812.
- Burr et al., "Gene mapping with recombinant inbreds in maize," *Genetics*, 1988, 118: 519-526.
- Burr et al., "Mapping Genes with Recombinant Inbreds," In Freeling and Walbot (Ed.), *The Maize Handbook*, (New York, Springer-Verlag, 1994), pp. 249-254.
- Bustos et al., "Regulation of b-glucuronidase expression in transgenic tobacco plants by an A/T-rich, cis-acting sequence found upstream of a French bean b-phaseolin gene" *Plant Cell*, 1(9):839-854 (1989).
- Carels et al., "Compositional properties of homologous coding sequences from plants" *J. Mol. Evol.*, 46:45-53 (1998).
- Cerdan et al., "A 146 bp fragment of the tobacco *Lhcb1\*2* promoter confers very-low-fluence, low-fluence and high-irradiance responses of phytochrom to a minimal CaMV 35S promoter" *Plant Mol. Biol.*, 33:245-255 (1997).
- Chan et al., "*Agrobacterium*-mediated production of transgenic rice plants expressing a chimeric a-amylase promoter/β-glucuronidase gene," *Plant Mol. Biol.*, 1993, 22: 491-506.
- Chen et al., "Functional analysis of regulatory elements in a plant embryo-specific gene" *Proc. Natl. Acad. Sci. USA*, 83:8560-8564 (1986).
- Cheng et al., "Highly divergent methyltransferases catalyze a conserved reaction in tocopherol and plastoquinone synthesis in cyanobacteria and photosynthetic eukaryotes," *Plant Cell*, 2003, 15: 2343-2356.
- Chenna et al., "Multiple sequence alignment with the Clustal series of programs" *Nucleic Acids Res.*, 31(13):3497-3500 (2003).
- Chickova et al., "Transgenic tobacco plants that overexpress alfalfa NADH-glutamate synthase have higher carbon and nitrogen content," *Journal of Experimental Botany*, 2001, 52(364): 2079-2087.
- Christou, "Strategies for variety-independent genetic transformation of important cereals, legumes and woody species utilizing particle bombardment" *Euphytica*, 85(1-3):13-27, (1995).
- Conceicao, "A cotyledon regulatory region is responsible for the different spatial expression patterns of *Arabidopsis* 2S albumin genes" *The Plant Journal*, 1994, 5(4):493-505.
- Conkling et al., "Isolation of transcriptionally regulated root-specific genes from tobacco" *Plant Physiol.*, 93:1203-1211, (1990).
- Conner and Domisse, "Monocotyledonous plants as hosts for *Agrobacterium*," *Int. J. Plant Sci.*, 1992, 153(4): 550-555.
- Dai et al., "RF2b, a rice bZIP transcription activator, interacts with RF2a and is involved in symptom development of rice tungro disease" *Proc. Natl. Acad. Sci. USA*, 101(2):687-692 (2004).
- de Feyter and Gaudron, *Methods in Molecular Biology*, vol. 74, Chapter 43, "Expressing Ribozymes in Plants", Edited by Turner, P.C., *Humana Press Inc.*, Totowa, NJ, 1997, pp. 403-415.
- deVicente and Tanksley, "QTL analysis of transgressive segregation in an interspecific tomato cross," *Genetics*, 1993, 134: 585-596.
- Dietrich et al., "AtPTR1, a plasma membrane peptide transporter expressed during seed germination and in vascular tissue of *Arabidopsis*," *Plant Journal*, 2004, 40(4): 488-499.
- Do et al., "ProbCons: Probabilistic consistency-based multiple sequence alignment," *Genome Res.*, 2005, 15(2): 330-340.
- Durbin et al., "3-Markov chains and hidden Markov models; 4-Pairwise alignment using HMMS; 5-Profile HMMs for sequence families" In *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*, (Cambridge University Press, Cambridge, 45 pages.
- Escudero et al., "T-DNA transfer in meristematic cells of maize provided with intracellular *Agrobacterium*" *Plant J.*, 10(2): 355-360 (1996).
- Evans et al., "Protoplasts Isolation and Culture in "Handbook of Plant Cell Culture," pp. 124-176, *MacMillan Publishing Company*, New York, 1983.
- Ezeagu et al., "Seed protein contents and nitrogen-to-protein conversion factors for some uncultivated tropical plant seeds," *Food Chemistry*, 2002, 78: 105-109.
- Fejes et al., "A 268 bp upstream sequence mediates the circadian clock-regulated transcription of the wheat Cab-1 gene in transgenic plants" *Plant Mol. Biol.*, 15:921-932 (1990).
- Fennoy et al., "Synonymous codon usage in *Zea mays* L. nuclear genes is varied by levels of C and G-ending codons" *Nucleic Acids Research*, 21(23):5294-5300 (1993).
- Foyer et al., "Adaptations of photosynthetic electron transport, carbon assimilation, and carbon partitioning in transgenic *Nicotiana plumbaginifolia* plants to changes in nitrate reductase activity," *Plant Physiology*, 1994, 104(1): 171-178.
- Fraisier et al., "Constitutive expression of a putative high-affinity nitrate transporter in *Nicotiana plumbaginifolia*: Evidence for post-transcriptional regulation by a reduced nitrogen source," *Plant Journal*, 2000, 23(4): 489-496.
- Fraley et al., "Genetic transformation in higher plants," *Crit. Rev. Plant. Sci.*, 1986, 4(1): 1-46.

- Frischauf et al., "Lambda replacement vectors carrying polylinker sequences," *J. Mol. Biol.*, 1983, 170: 827-842.
- Fromm et al., "An octopine synthase enhancer element directs tissue-specific expression and binds ASF-1, a factor from tobacco nuclear extracts" *The Plant Cell*, 1:977-984 (1989).
- Fromm et al., "Inheritance and expression of chimeric genes in the progeny of transgenic maize plants," *Biotechnology*, 1990, 8: 833-844.
- Gardiner et al., "Development of a core RFLP map in maize using an immortalized F2 population," *Genetics*, 1993, 134: 917-930.
- GenBank Accession No. AAC28509, dated Aug. 4, 1998, 2 pages.
- GenBank Accession No. AAO24547, dated Jan. 21, 2003, 3 pages.
- GenBank Accession No. AF096096, dated Jan. 25, 1999, 2 pages.
- GenBank Accession No. AF129516, dated Apr. 6, 1999, 2 pages.
- GenBank Accession No. AL163815.1, dated Nov. 14, 2006, 49 pages.
- GenBank Accession No. BAC43284, dated Feb. 14, 2004, 2 pages.
- GenBank Accession No. CAB87794, dated Nov. 15, 2006, 2 pages.
- GenBank Accession No. CAB87717, dated Nov. 14, 2006, 3 pages.
- GenBank Accession No. L05934, dated Oct. 22, 1993, 3 pages.
- GenBank Accession No. NM\_112313, dated Jun. 9, 2006, 3 pages.
- GenBank Accession No. NM\_121195, dated Jan. 10, 2002, 3 pages.
- GenBank Accession No. NM\_101234, dated Jun. 9, 2006, 3 pages.
- GenBank Accession No. NM\_103793, dated Jun. 9, 2006, 2 pages.
- GenBank Accession No. NM\_122025, dated Jun. 9, 2006, 2 pages.
- GenBank Accession No. NM\_130084, dated Jun. 9, 2006, 3 pages.
- GenBank Accession No. NM\_180224, dated Jun. 9, 2006, 3 pages.
- GenBank Accession No. NP\_175330, dated Jun. 9, 2006, 2 pages.
- GenBank Accession No. NP\_182046, dated Jun. 9, 2006, 2 pages.
- GenBank Accession No. NP\_188071, dated Jun. 9, 2006, 3 pages.
- GenBank Accession No. NP\_196718, dated Jun. 9, 2006, 2 pages.
- GenBank Accession No. NP\_197518, dated Jun. 9, 2006, 2 pages.
- GenBank Accession No. NP\_563930, dated Jun. 9, 2006, 3 pages.
- GenBank Accession No. NP\_850555, dated Jun. 9, 2006, 2 pages.
- GenBank Accession No. U93215, dated Feb. 27, 2002, 42 pages.
- GenBank Accession No. Q9LPV5, dated Mar. 1, 2004, 3 pages.
- Ghosh et al., "Transgenic Indica rice (*Oryza sativa* L.) plants obtained by direct gene transfer to protoplasts" *J. Biotechnol.*, 32:1-10 (1994).
- Gleave, AP., "A versatile binary vector system with a T-DNA organizational structure conducive to efficient integration of cloned DNA into the plant genome" *Plant Mol. Biol.* 20:1203-1207 (1992).
- Gordon-Kamm et al., "Transformation of maize cells and regeneration of fertile transgenic plants," *Plant Cell*, 1990, 2: 603-618.
- Guo et al., "Protein tolerance to random amino acid change," *PNAS*, 2004, 101(25): 9205-9210.
- Gould et al., "Transformation of *Zea mays* L. Using *Agrobacterium tumefaciens* and the shoot apex" *Plant Physiology*, 95:426-434 (1991).
- Graves and Goldman, "The transformation of *Zea mays* seedling with *Agrobacterium tumefaciens*" *Plant Mol. Biol.*, 7:43-50 (1986).
- Green et al., "Binding site requirements for pea nuclear protein factor GT-1 correlate with sequences required for light-dependent transcriptional activation of the *rbcS-3A* gene" *EMBO J.*, 7:4035-4044 (1988).
- Hamilton et al., "Stable transfer of intact high molecular weight DNA into plant chromosomes," *Proc. Natl. Acad. Sci., USA*, 1996, 93: 9975-9979.
- Hamilton, "A binary-BAC system for plant transformation with high-molecular-weight DNA" *Gene*, 200:107-116 (1997).
- Heath et al., "Analysis of storage proteins in normal and aborted seeds from embryo-lethal mutants of *Arabidopsis thaliana*," *Planta*, 1986, 169: 304-312.
- Henkel, "Soy: Health claims for soy protein, questions about other components," *FDA Consumer*, 2000, 3 pages.
- Herrera-Estrella et al., "Chimeric genes as dominant selectable markers in plant cells" *EMBO J.*, 2(6):987-995 (1983).
- Hiei et al., "Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA," *Plant Journal*, 1994, 6(2): 271-282.
- Hill et al., "Carbon supply for storage-product synthesis in developing seeds of oilseed rape," *Biochemical Society Transactions*, 2000, 28(6): 667-669.
- Hong et al., "Promoter sequences from two different *Brassica napus* tapetal oleosin-like genes direct tapetal expression of  $\beta$ -glucuronidase in transgenic *Brassica* plants" *Plant Mol Biol.*, 1997 34(3):549-555.
- Hosoyama et al. "Oryzacystatin exogenously introduced into protoplasts and regeneration of transgenic rice" *Biosci. Biotechnol. Biochem.* 58(8): 1500-1505 (1994).
- Huang et al., "Cloning and functional characterization of an *Arabidopsis* nitrate transporter gene that encodes a constitutive component of low-affinity uptake," *Plant Cell*, 1999, 11: 1381-1392.
- Huynh et al., In Glover NM (ed) DNA Cloning: A practical approach, vol. 1 (Oxford, IRL Press, 1985) , pp. 49-78.
- Hwang et al., "Aleurone- and embryo-specific expression of the  $\beta$ -glucuronidase gene controlled by the barley *Chi26* and *Ltp1* promoters in transgenic rice" *Plant Cell Rep.* 20(7):647-654 (2001).
- Hwang et al., "Cloning and functional characterization of an *Arabidopsis* nitrate transporter gene that encodes a constitutive component of low-affinity uptake," *Plant Cell*, 1999, 11(8): 1381-1392.
- Hyrup et al., Peptide nucleic acids (PNA): synthesis, properties and potential applications *Bioorgan. Med. Chem.*, 4:5-23 (1996).
- Ishida et al., "High efficiency transformation of maize (*Zea mays* L. mediated by *Agrobacterium tumefaciens*" *Nature Biotechnology*, 14:745-750 (1996).
- Jako et al., "Seed-specific over-expression of an *Arabidopsis* cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight," *Plant Physiology*, 2001, 126(2): 861-874.
- Jordano et al., "A sunflower helianthinin gene upstream sequence ensemble contains an enhancer and sites of nuclear protein interaction" *Plant Cell*, 1:855-866 (1989).
- Katavic et al., "Utility of the *Arabidopsis*FAE1 and yeast SLC1-1 genes for improvements in erucic acid and oil content in rapeseed," *Biochemical Society Transactions*, 2000, 28(6): 935-937.
- Keller and Baumgartner, "Vascular-specific expression of the bean GRP 1.8 gene is negatively regulated" *Plant Cell*, 3(10):1051-1061 (1991).
- Keller and Manak "DNA Probes Section One: Molecular Hybridization Technology", 2nd Ed. c. 1993 by Stockton Press, New York, NY, pp. 1-25.
- Keskin et al., "A new, structurally nonredundant, diverse data set of protein-protein interfaces and its implications," *Protein Science*, 2004, 13: 1043-1055.
- Klee et al. "Agrobacterium-mediated plant transformation and its further applications to plant biology" *Ann. Rev. of Plant Phys.*, 38:467-486 (1987).
- Klein et al., "High-velocity microprojectiles for delivering nucleic acids into living cells," *Nature*, 1987, 327: 70-73.
- Koch et al., "Reduced amino acid content in transgenic potato tubers due to antisense inhibition of the leaf H<sup>+</sup>/amino acid symporter StAAP1," *Plant Journal*, 2003, 33(2): 211-220.
- Komari et al., "Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by *Agrobacterium tumefaciens* and segregation of transformants free from selection markers" *Plant J.*, 10(1): 165-174 (1996).
- Kozziel et al., "Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*," *Biotechnology*, 1993, 11: 194-200.
- Lam et al., "Site-specific mutations in alter in vitro factor binding and change promoter expression pattern in transgenic plants" *Proc. Natl. Acad. Sci. USA*, 86:7890-7894 (1989).
- Lassner et al., "A jojoba beta-Ketoacyl-CoA synthase cDNA complements the canola fatty acid elongation mutation in transgenic plants," *Plant Cell*, 1996, 8(2): 281-292.
- Li et al., "Oil content of arabidopsis seeds: The influence of seed anatomy, light and plant-to-plant variation," *Phytochemistry*, 2006, 67: 904-915.
- Liljegren, "Interactions among *APETALA1*, *LEAFY*, and *Terminal FLOWER1* specify meristem fate" *Plant Cell*, 11:1007-1018 (1999).
- Luan et al., "A rice *cab* gene promoter contains separate cis-acting elements that regulate expression in dicot and monocot plants" *The Plant Cell*, 4:971-981 (1992).
- Lubberstedt et al., "Promoters from genes for plastid proteins possess regions with different sensitivities toward red and blue light" *Plant Physiol.*, 104:997-1006 (1994).

- Marra et al., "High throughput fingerprint analysis of large-insert clones," *Genomic Research*, 1997, 7: 1072-1084.
- Matsuoka et al., "Tissue-specific light-regulated expression directed by the promoter of a C4 gene, maize pyruvate, orthophosphate dikinase, in a C3 plant, rice" *Proc. Natl. Acad. Sci. USA*, 90:9586-9590 (1993).
- Matteucci et al. "Synthesis of Deoxyoligonucleotides on a polymer support" *J. Am. Chem. Soc.*, 103:3185-3191 (1981).
- May et al., "Generation of transgenic banana (*Musa acuminata*) plants via *Agrobacterium*-mediated transformation" *Bio/Technology*, 13:486-492 (1995).
- McAlister-Henn et al., "Application of the yeast two-hybrid system," *Methods*, 1999, 19: 330-337.
- McClure et al., "Transcription, Organization, and Sequence of an Auxin-Regulated Gene Cluster in Soybean," *Plant Cell*, 1989, 1:229-239.
- McCormac et al., "A flexible series of binary vectors for agrobacterium-mediated plant transformation" *Mol. Biotechnol.*, 8:199-213 (1997).
- McCormick et al., "Leaf disc transformation of cultivated tomato (*L. esculentum*) using *Agrobacterium tumefaciens*," *Plant Cell Reports*, 1986, 5: 81-84.
- Medberry et al., "The Commelina yellow mottle virus promoter is a strong promoter in vascular and reproductive tissues" *Plant Cell*, 4(2):185-192 (1992).
- Meier et al., "Elicitor-inducible and constitutive in vivo DNA footprints indicate novel *cis*-acting elements in the promoter of a parsley gene encoding pathogenesis-related protein 1" *Plant Cell*, 3:309-316 (1991).
- Müller et al., "High meiotic stability of a foreign gene introduced into tobacco by *Agrobacterium*-mediated transformation" *Mol. Gen. Genet.*, 207:171-175 (1987).
- Needleman and Wunsch "A general method applicable to the search for similarities in the amino acid sequence of two proteins" *J. Mol. Biol.*, 48:443-453 (1970).
- Nehra et al., "Self-fertile transgenic wheat plants regenerated from isolated scutellar tissues following microprojectile bombardment with two distinct gene constructs," *Plant Journal*, 1994, 5(2): 285-297.
- Panaud et al., "Frequency of microsatellite sequences in rice (*Oryza sativa* L.)," *Genome*, 1995, 38:1170-1176.
- Paszkowski et al. "Direct gene transfer to plants" *EMBO J.*, 3:2717-2722 (1984).
- Pearson and Lipman "Improved tools for biological sequence comparison" *Proc. Natl. Acad. Sci. USA*, 85: 2444-2448 (1988).
- Perriman et al., "Effective ribozyme delivery in plant cells" *Proc. Natl. Acad. Sci. USA*, 92(13):6175-6179 (1995).
- Pfam web cite (describing consensus sequences for a variety of protein motifs and domains at sanger.ac.uk/Pham and genome.wustl.edu.Pfan, 2006, 2 pages.
- Prioli and Soendahl, "Plant regeneration and recovery of fertile plants from protoplasts of maize (*Zea mays* L.)," *Bio/Technology*, 1989, 7: 589-594.
- Refseth et al., "Hybridization capture of microsatellites directly from genomic DNA" *Electrophoresis*, 18:1519-1523 (1997).
- Rhee, "Determination of total nitrogen" in Handbook of Food Analytical Chemistry—Water, Proteins, Enzymes, Lipids, and Carbohydrates. (R. Wrolstad et al, ed.) *John Wiley and Sons, Inc.*, 2005, p. 105-113.
- Ritala et al., "Fertile transgenic barley by particle bombardment of immature embryos," *Plant Mol. Biol.*, 1994, 24: 317-325.
- Riggs et al., "Cotyledon nuclear proteins bind to DNA fragments harboring regulatory elements of phytohemagglutinin genes" *Plant Cell*, 1(6):609-621 (1989).
- Ritchie et al., "*Agrobacterium tumefaciens*-mediated expression of *gusA* in maize tissues," *Transgenic Res.*, 1993, 2: 252-265.
- Rivera et al., "Genomic evidence for two functionally distinct gene classes" *Proc. Natl. Acad. Sci. USA*, 95:6239-6244 (1998).
- Salomon et al., "Genetic identification of functions of TR-DNA transcripts in octopine crown galls" *EMBO J.*, 3:141-146 (1984).
- Sambrook et al., 1989, "Molecular Cloning, A Laboratory Manual", second edition, *Cold Spring Harbor Press*, Plainview; NY, 21 pages.
- Senior et al., "Simple sequence repeat markers developed from Maize sequences found in the Genbank database: Map construction," *Crop Science*, 1996, 36:1676-1683.
- Shen et al., "The homeobox gene *GLABRA2* affects seed oil content in *Arabidopsis*," *Plant Molecular Biology*, 2006, 60 (3): 377-387.
- Smith and Waterman, "Comparison of Biosequences" *Advances in Applied Mathematics.*, 2:482-489 (1981).
- Sheridan, "The *macl* Gene: Controlling the commitment to the meiotic pathway in Maize" *Genetics*, 142:1009-1020 (1996).
- Shillito et al., "Regeneration of fertile plants from protoplasts of elite inbred maize," *Bio/Technology*, 1989, 7: 581-587.
- Shizuya et al., "Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector," *Proc. Natl. Acad. Sci. USA*, 1992, 89: 8794-8797.
- Singh et al., "Transgenic expression of a delta 12-epoxygenase gene in *Arabidopsis* seeds inhibits accumulation of linoleic acid," *Planta*, 2001, 212: 872-879.
- Slocombe et al., "Temporal and tissue-specific regulation of a *Brassica napus* stearoyl-acyl carrier protein desaturase gene" *Plant Physiol.*, 104(4):1167-1176 (1994).
- Sonnhammer et al., "Pfam: A comprehensive database of protein domain families based on seed alignments" *Proteins*, 28:405-420 (1997).
- Sonnhammer et al., "Pfam: multiple sequence alignments and HMM-profiles of protein domains" *Nucl. Acids Res.*, 26:320-322 (1998).
- Spencer et al., "Bialaphos selection of stable transformants from maize cell culture," *Theor. Appl. Genet.*, 1990, 79: 625-631.
- Sternberg et al., "Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs," *Proc. Natl. Acad. Sci. USA*, 1990, 87(1): 103-107.
- Summerton and Weller, "Morpholino antisense oligomers: design, preparation, and properties" *Antisense Nucleic Acid Drug Dev.*, 7:187-195 (1997).
- Tanksley and McCouch, "Seed banks and molecular maps: Unlocking genetic potential from the wild," *Science*, 1997, 277: 1063-1066.
- Taramino et al., "Simple sequence repeats for germplasm analysis and mapping in maize," *Genome*, 1996, 39(2): 277-287.
- Thornton et al., "From structure to function: approaches and limitations," *Nature Structural Biology*, 2000, pp. 991-994.
- Tijessen, "Hybridization with Nucleic Acid Probes" in Laboratory Techniques in Biochemistry and Molecular Biology, P.C. van der Vliet, ed., c. 1993 by Elsevier, Amsterdam, pp. 19-78.
- Tomlinson et al., "Evidence that the hexose-to sucrose ration does not control the switch to storage product accumulation in oilseeds: analysis of tobacco seed development and effects of overexpressing apoplatic invertase," *Journal of Experimental Botany*, 2004, 55(406): 2291-2303.
- Truernit et al., "The promoter of the *Arabidopsis thaliana* *SUC2* sucrose-H<sup>+</sup> symporter gene directs expression of  $\beta$ -glucuronidase to the phloem: Evidence for phloem loading and unloading by *SUC2*" *Planta*. 196:564-570 (1995).
- Tsay et al., "Nitrate transporters and peptide transporters," *FEBS Letters*, 2007, 581: 2290-2300.
- Tuskan et al., "The genome of black cottonwood, *Populus trichocarpa*," *Science*, 2006, 313(5793): 1596-1604.
- Urao et al. "Molecular cloning and characterization of a gene that encodes a MYC-related protein in *Arabidopsis*" *Plant Mol. Biol.*, 32:571-576 (1996).
- Urdea et al. "Chemical synthesis of a gene for human epidermal growth factor urogastrone and its expression in yeast" *Proc. Natl. Acad. Sci. USA*, 80:7461-7465 (1983).
- Vasil et al., "Rapid production of transgenic wheat plants by direct bombardment of cultured immature embryos," *Bio/Technology*, 1993, 11: 1553-1558.
- Venkateswarlu et al., "Evidence for T-DNA mediated gene targeting to tobacco chloroplasts" *Biotechnology*, 9:1103-1105 (1991).
- Walden and Cowan, "A novel 205-kilodalton testis-specific serine/threonine protein kinase associated with microtubules of the spermatid manchette," *Mol. Cell Biol.*, 1993, 13(12): 7625-7635.
- Wan and Lemaux, "Generation of large number of independently transformed fertile barley plants," *Plant Physiol.*, 1994, 104: 37-48.

- Weber et al., "Interaction of cytosolic and plastidic nitrogen metabolism in plants," *Journal of Experimental Botany*, 2002, 53(370): 865-874.
- Weising et al., "Foreign genes in plants: transfer, structure, expression, and applications" *Ann. Rev. Genet.*, 22:421-477 (1988).
- Willmitzer (Willmitzer, L. (1993) Transgenic plants. In : *Biotechnology, A Multi-Volume Comprehensive Treatise* (H.G. Rehm, G. Reed, A. Puehler, P. Stadler, eds.). vol. 2, 627-659, VCH Weinheim-New York-Basel-Cambridge).
- Wilmink et al., "Expression of the GUS-gene in the monocot tulip after introduction by particle bombardment and *Agrobacterium*," *Plant Cell Reports*, 1992, 11: 76-80.
- Yamamoto et al., "The promoter of a pine photosynthetic gene allows expression of a  $\beta$ -glucuronidase reporter gene in transgenic rice plants in a light-independent but tissue-specific manner" *Plant Cell Physiol.*, 1994, 35:773-778.
- Yan et al., "New construct approaches for efficient gene silencing in plants," *Plant Physiology*, 2006, 141: 1508-1518.
- Yanagisawa et al., "Metabolic engineering with Dof1 transcription factor in plants: Improved nitrogen assimilation and growth under low-nitrogen conditions" *Proc. Natl. Acad. Sci. USA*, 101(20):7833-7838 (2004).
- Zhang et al., "DNA sequences that activate isocitrate lyase gene expression during late embryogenesis and during postgerminative growth," *Plant Physiology*, 110:1069-1079 (1996).
- Zheng et al., "SPK1 is an Essential S-Phase-Specific Gene of *Saccharomyces cerevisiae* that encodes a nuclear serine/threonine/tyrosine kinase" *Mol. Cell Biol.*, 1993, 13:5829-5842.

\* cited by examiner

gi 2739168	Lead:Lead117-CeresClone41573	-----MS	LKINLLVCLF	I FHLVAASV	RGH-----K	IGIYEIKKGD	36
gi 50920801	CeresClone:1371577	MAAT SARSA	WQWAAVAAAV	AAASLGGTTP	AEARRGCKES	IGFYELRRGE	0
CeresClone:1560908	CeresClone:1314177	-----MARA	LLVALLLVSC	FVALGSCRA	DA--AERKIT	MGVYELNKGD	50
Consensus		-----MARA	PLLLALLCAC	AAAAGSANA	AGR-----KM	MGVYELKAGD	42
		-----MA-A	PLLLLLAAVC	LACFAPGANA	AGR-----KM	MGVYELSKGD	39
			-LLLLLL-AC	-AA--A--A	AG-----K-	VGVEYELKKGD	39
							50
gi 2739168	Lead:Lead117-CeresClone41573	FSVKITNYGT	SII SVLLPDK	HGKI GDVVLG	YDTIEEYKND	TSYFGATLGR	86
gi 50920801	CeresClone:1371577	FSVVTNWGA	TI LTVNLPDK	MDI VLG	YDSVKTYKTD	KVYFGATVGR	27
CeresClone:1560908	CeresClone:1314177	FSIKVTNWGA	TLMSVILLDS	NGHI DDVVLG	YKDJ GSYVND	TTYFGALVGR	100
Consensus		FSVGVTNWGA	TTTSVVLPDS	RCNFGDVLG	YDTVAEYMNG	TAYFGGLIGR	92
		FSAKVTNWGA	TVTSVVFPPDS	KGDFADVVLG	YDTI GGYVSG	KSYFGALVGR	89
			TI-SVILPDS	KGNI GDVVLG	YDTIGEYVNG	SSYFGALVGR	89
							100
gi 2739168	Lead:Lead117-CeresClone41573	VANRI GGAQF	LLNGLHYKELV	PNEGKMLHG	GPKCF SKVVW	KMSKYVKDGP	136
gi 50920801	CeresClone:1371577	VANRI GKGKF	KLNGKEYKTS	VNDGKNTLHG	GKKGFGDMVW	AMAKHOYDCK	77
CeresClone:1560908	CeresClone:1314177	VANRI AGGRF	TVKGFHAFHITY	KNDGKNTLHG	CHRGFNQVFM	SVRERAITGH	149
Consensus		VANRI ANARF	LLDGKITVRLV	PNDGNNTLHG	CHRGFQKVVW	TVKEHVAAGS	142
		VANRI ANARF	VLDGKAYHLF	KNDGNNTLHG	CHRGF SQVVW	TVKEFVGGGD	139
		VANRVAARF	VLDGKMYHLY	ANDGKNALHG	CHRGF SKVI M	TVKEYVSGGD	139
			TLDGK-Y-LY	-NDGKNTLHG	CHRGF SKVVW	TVKE-V-GG-	150
gi 2739168	Lead:Lead117-CeresClone41573	CPYITLIYYYS	ADGEEGFPGA	VLASVTYITLK	DSYKLSVMFR	AKALN--KAT	184
gi 50920801	CeresClone:1371577	KPHIVFTHTS	PDCDQGF PGE	LSMTVTYKLV	KDNELSVMME	AKPKD--KAT	125
CeresClone:1560908	CeresClone:1314177	FPYITFYIRS	YDGEQGFPGA	LQVVTYKID	GDYSFSVTMY	ARPVDGKPT	199
Consensus		SPFITLYYHS	FDGEQGFPGD	VDVHVTYQLS	SPHVL SVR MN	ATARG--KAT	190
		SPYITLYYHS	FDGEQGFPGD	LQVVTYQLA	APYVLRHMN	ATAVN--KAT	187
		SPHITLYYHS	FDGEQGFPGA	LQVVTYELS	SPYVLSVRMN	ATALD--KAT	187
			FDGEQGFPG-	LQVVTY-LS	SPYVLSV-MN	A-A-D--KAT	200

Figure 1

gi 2739168	NI	GG	HD	SG	DV	LS	QI	FG	SS	HI	L	V	D	K	Q	L	I	PT	GE	I	A	P	I	K	N	234						
Lead:Lead117-CeresClone41573	PV	NL	AH	HS	Y	W	L	S	E	E	Q	I	L	G	S			P	T	C	K	I	N	P	V	K	G	175				
gi 50920801	PV	NL	AQ	H	T	Y	W	L	D	H	S	V	Q	I	F	A	S	A	V	T	P	V	G	A	G	L	I	249				
CeresClone:1371577	PV	NL	AH	H	A	Y	W	L	G	E	E	V	R	L	F	A	S	R	Y	T	P	V	D	A	S	L	I	240				
CeresClone:1560908	PV	NL	AQ	H	T	Y	W	L	R	N	T	V	Q	L	F	A	S	R	Y	T	P	V	G	G	L	I	I	237				
CeresClone:1314177	PV	NL	AQ	H	T	Y	W	L	N	G	N	T	V	Q	L	F	A	A	R	Y	T	P	V	D	A	I	L	I	237			
Consensus	PV	N	L	A	-	H	T	Y	W	L	-	T	V	Q	-	F	A	S	R	Y	T	P	V	D	A	-	L	I	250			
gi 2739168	T	P	Y	D	F	L	K	P	R	K	V	G	S	R	I	N	K	L	K	N	Y	D	I	N	Y	V	L	D	S	T	272	
Lead:Lead117-CeresClone41573	T	A	Y	D	F	L	Q	L	R	P	I	K	D	N	M	K	I	L	K	T	Y	D	I	N	Y	V	L	D	S	T	214	
gi 50920801	T	P	F	D	F	R	A	P	A	P	P	G	A	R	I	A	D	V	P	G	Y	D	I	N	Y	V	L	D	S	T	291	
CeresClone:1371577	T	P	Y	D	F	R	P	A	A	V	G	S	R	I	G	G	L	S		Y	D	I	N	Y	V	L	D	S	T	281		
CeresClone:1560908	T	P	Y	D	F	L	A	P	A	A	V	G	S	R	I	R	Q	V	S	G	Y	D	I	N	Y	V	L	D	S	T	282	
CeresClone:1314177	T	P	Y	D	F	R	A	P	A	A	V	R	E	H	L	R	Q	V	V	G	Y	D	I	N	Y	V	L	D	S	T	285	
Consensus	T	P	Y	D	F	-	A	P	A	A	V	G	S	R	I	R	Q	-	-	-	Y	D	I	N	Y	V	-	D	S	T	300	
gi 2739168	V	V	R	D	G	A	S	G	R	A	M	D	V	Q	A	S	S	P	G	V	T	K	G	K	G	F	V	Y	Q		322	
Lead:Lead117-CeresClone41573	E	L	V	D	K	K	S	C	R	K	M	E	L	S	G	N	Q	A	G	L	V	K	G	K	N	G	A	A	Y	Q	264	
gi 50920801	V	V	S	E	P	T	S	G	R	V	L	E	L	W	S	D	Q	P	G	L	D	E	G	K	G	A	R	M	V		341	
CeresClone:1371577	V	V	R	D	G	A	S	G	R	A	M	E	L	W	A	D	Q	P	G	V	V	R	G	K	G	K	V	Y	G		331	
CeresClone:1560908	V	V	R	D	G	A	S	G	R	A	L	E	L	W	A	N	Q	P	G	V	V	K	G	K	G	K	V	Y	Q		332	
CeresClone:1314177	V	V	R	D	G	A	S	G	R	A	L	E	L	W	A	D	Q	P	G	V	V	K	G	K	G	S	V	Y	G		335	
Consensus	V	V	R	D	G	A	S	G	R	A	-	E	L	W	A	D	Q	P	G	V	Q	F	Y	T	G	N	F	L	-	D	350	
gi 2739168	V	F	P	D	A	V	N	H	P	N	F	P	S	Q	I	V	R	P	G	V	F	S	I	K	K	-					356	
Lead:Lead117-CeresClone41573	S	Y	P	D	A	L	N	H	P	K	F	P	S	Q	I	V	E	P	G	K	F	S	I	V	S	-					298	
gi 50920801	D	Y	P	D	A	V	N	H	N	A	K	F	P	T	E	I	Y	R	K	G	Q	F	S	L	A	K						376
CeresClone:1371577	G	F	P	D	A	V	N	H	P	S	F	P	S	Q	I	V	R	P	G	K	F	S	F	-	-						363	
CeresClone:1560908	G	F	P	D	A	V	N	H	P	D	F	P	S	Q	I	V	R	P	G	Q	F	S	F	-	-						364	
CeresClone:1314177	G	F	P	D	A	V	N	H	P	E	F	P	S	Q	I	V	R	P	G	V	F	S	F	-	-						368	
Consensus	G	F	P	D	A	V	N	H	P	-	F	P	S	Q	I	V	R	P	G	Q	-	K	Y	K	H	D	M	-	F	K	386	

Figure 2

CeresClone:953928	MKVVAAFLLA	VL	SG	KA	C	P	T	S	AD	I	K	M	L	N	S	V	G	C	E	I	E	D	S	Q	I	E	L	L	K	E	V	N	G	K	D	50		
CeresClone:426736	MKVVAAAYLLA	VL	SG	KA	S	P	T	S	AD	I	K	T	L	G	S	V																				50		
Lead-Lead118-CeresClone25429	MKVAAFLLA	VL	GG	N	A	N	P	S	E	N	I	K	D	I	G	A	V																			50		
CeresClone:524682	MKVVAAYLLA	VL	GG	N	S	P	S	A	D	V	K	E	L	G	S	V																				50		
CeresGdna:1488311	MKVVAAYLLA	VL	GG	N	C	P	T	A	E	D	L	K	H	L	G	S	V																			50		
CeresClone:1609735	MKVVAAYLLA	LL	GG	N	S	P	S	A	E	D	L	K	N	L	G	S	V																			50		
gi 42565379	MKVVAAYLLA	VL	GG	N	T	P	S	A	E	V	K	D	I	L	G	S	V																			50		
gi 24473796	MKVVAAYLLA	VL	GG	N	T	P	S	A	E	D	L	K	D	I	L	G	S	V																		50		
Consensus	MKVVAAYLLA	VL	GG	N	T	S	P	S	A	E	D	I	K	D	L	G	S	V	GA	--	DD	R	I	E	L	L	S	E	V	K	G	K	D	50				
CeresClone:953928	VAELI	A	M	G	R	E	K	L	A	S	V	P	S	-	G	G	G	G	G	G	G	G	G	A	A	P	-	A	E	D	K	K	E	E	K	E	98	
CeresClone:426736	LAELI	A	A	A	G	R	E	K	L	A	S	V	P	S	-	G	G	G	G	G	G	G	G	A	A	P	-	A	E	S	K	K	E	E	K	E	100	
Lead-Lead118-CeresClone25429	VAELI	A	A	S	G	R	E	K	L	A	S	V	P	S	-	G	G	G	A	A	A	A	P	-	A	E	H	-	A	E	K	K	E	A	K	E	98	
CeresClone:524682	VAELI	A	A	A	G	R	E	K	L	A	T	V	P	S	-	G	C	G	A	A	A	A	P	-	A	E	A	A	A	K	K	E	E	K	V	E	99	
CeresGdna:1488311	TELI	A	S	G	R	E	K	L	A	S	V	P	S	-	G	G	A	A	A	A	A	P	-	A	E	A	A	A	A	A	K	K	E	E	K	V	E	99
CeresClone:1609735	TELI	A	S	G	R	E	K	L	A	S	V	P	S	-	G	G	A	A	A	A	A	P	-	A	E	S	G	A	A	A	A	A	A	A	A	A	A	99
gi 42565379	TELI	A	S	G	R	E	K	F	A	S	V	P	S	-	G	G	A	A	P	-	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	99	
gi 24473796	TELI	A	S	G	R	E	K	L	A	S	V	P	S	-	G	G	A	A	P	-	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	99	
Consensus	I	T	E	L	I	A	S	G	R	E	K	L	A	S	V	P	S	-	G	G	G	A	A	-	A	A	-	A	A	E	-	K	K	E	E	K	V	100
CeresClone:953928	EKEES	D	D	D	D	M	G	F	S	L	F	E	113																									
CeresClone:426736	EKEES	D	D	D	D	M	G	F	S	L	F	E	115																									
Lead-Lead118-CeresClone25429	EKEES	D	D	D	D	M	G	F	S	L	F	E	113																									
CeresClone:524682	EKEES	D	D	D	D	M	G	F	S	L	F	D	114																									
CeresGdna:1488311	EKEES	D	D	D	D	M	G	F	S	L	F	D	114																									
CeresClone:1609735	R	100																																				
gi 42565379	EKEES	D	E	D	D	M	G	F	S	L	F	D	114																									
gi 24473796	EKED	T	D	D	D	M	G	F	S	L	F	D	113																									
Consensus	EKEES	D	D	D	D	M	G	F	S	L	F	D	115																									

Figure 2 (continued)



gi 55978016	MPQI QYSEKY	YDDVYEYRHH	VLPPDI AKLL	PKNRLLTETE	WRGLGVQQSR	50
CeresClone:947192	MGQI QYSDKY	FDDT FEYRHH	VLPPDI AKLL	PKNRLLSENE	WRAI GVQQSR	50
gi 27435806	MGQI QYSEKY	LDDT FEYRHH	VLPEVAKLL	LKNRLLSENE	WRAI GVQQSR	50
gi 45935118	MGQI QYSEKY	FDDT FEYRHH	VLPEVAKLL	PKNRLLSENE	WRAI GVQQSR	50
Lead·Lead119·CeresClone5750	MGQI QYSEKY	FDDT FEYRHH	VLPEVAKLL	PKNRLLSENE	WRAI GVQQSR	50
gi 42362268	MGQI QYSEKY	FDDT FEYRHH	VLPEVAKLL	PKNRLLSENE	WRAL GVQQSR	50
CeresClone:1017141	MGQI QYSEKY	FDDT FEYRHH	VLPEVAKLL	PKNRLLAENE	WRAL GVQQSR	50
CeresClone:1448636	MGQI QYSEKY	FDDT FEYRHH	VLPEVAKLL	PKNRLLAENE	WRAL GVQQSR	50
gi 50919707	MGQI QYSEKY	FDDT FEYRHH	VLPEVAKLL	PKNRLLSENE	WRAI GVQQSR	50
Consensus	MGQI QYSEKY	FDDT FEYRHH	VLPEVAKLL	PKNRLLSENE	WRAI GVQQSR	50
gi 55978016	GWVHYAI HRP	EPHI MLYRRP	LTETPPQ	HAQN	77	
CeresClone:947192	GWVHYAI HRP	EPHI MLFRRP	LN YQQQQEN	VLAK	88	
gi 27435806	GWVHYAI HRP	EPHI MLFRRP	LN YQQQQDQA	LVK	83	
gi 45935118	GWVHYAI HRP	EPHI MLFRRP	LN YQQQESQ	LIAK	88	
Lead·Lead119·CeresClone5750	GWVHYAVHRP	EPHI MLFRRP	LN YQQQENQ	MLVK	87	
gi 42362268	GWVHYAI HRP	EPHI MLFRRP	LN YQQQENQ	MLVK	88	
CeresClone:1017141	GWVHYAVHRP	EPHI MLFRRP	LN YQQQDAA	AAQAQMPK	90	
CeresClone:1448636	GWVHYAVHRP	EPHI MLFRRP	LN YQQQDAA	AAQAQMPK	90	
gi 50919707	GWVHYAI HRP	EPHI MLFRRP	LN FQQQEA	AAQAQMLPK	90	
Consensus	GWVHYAI HRP	EPHI MLFRRP	LN YQQQE	A-----ML-K	90	

Figure 3

CeresGdna:1440705	MTTQQGPPAA	SYSPARDGGG	GRSAEDEEKE	KAAAI DNWLP	I SATRNAKWW	0
CeresClone:1571117	---MGJQAP	ENYPAEKVQD	ARSAREK---	---AI DDWLP	TSSRNAKWW	50
Lead:Lead120-CeresClone218626	---	---MAKQWQD	GRSAQEK---	---AI DDWLP	TSSRNAKWW	40
gi 30409136	---	---	---	---	---	32
Consensus	---G--A-	---PA---QD	GRSA-EK---	---AI DDWLP	ITSSRNAKWW	50
CeresGdna:1440705	YSAFHNV TAM	VGAGVLSL PY	AMANL GWGPG	TVLVL SWI	LYTLWQMVE	41
CeresClone:1571117	YSAFHNV TAM	VGAGVLSL PY	AMSQLGWAG	MLLSWVI	LYTLWQMVE	100
Lead:Lead120-CeresClone218626	YSAFHNV TAM	VGAGVLSL PY	AMSELGWGPG	JAVLVVSWVI	LYTLWQMVE	90
gi 30409136	YSAFHNV TAM	VGAGVLSL PY	AMSELGWGPG	JAVLILSWII	LYTLWQMVE	82
Consensus	YSAFHNV TAM	VGAGVLSL PY	AMSELGWGPG	I A-LVLSWVI	LYTLWQMVE	100
CeresGdna:1440705	MHEMVP GKRF	DRYHEL GQHA	FGEKLG L Y I V	VPQQLTCEVG	VDI VYMTGG	91
CeresClone:1571117	MHEMVP GKRF	DRYHEL GQHA	FGDRLGLWIV	VPQQLVVEVG	VNI VYMTGG	150
Lead:Lead120-CeresClone218626	MHEMVP GKRF	DRYHEL GQHA	FGDRLGLWIV	VPQQLVVEVG	VNI VYMTGG	140
gi 30409136	MHEMVP GKRF	DRYHEL GQHA	FGEKLG L W I V	VPQQLVVEVG	VNI VYMTGG	132
Consensus	MHEMVP GKRF	DRYHEL GQHA	FG--LGLWIV	VPQQLVVEVG	VNI VYMTGG	150
CeresGdna:1440705	KSLQKFNH L	VCKD	CAPI KLT YFI	MI FASVHFV L	SHLPNFNSI S	134
CeresClone:1571117	TSLQKFHDV L	VCGDAA-CEG	GRK RLTYFI	MI FASVHFV L	AQLPNFDSI S	199
Lead:Lead120-CeresClone218626	KSLKKFHDV L	VCGDITGVCEG	KDNI KLT YFI	MI FASVHFV L	SQLPNFNSI S	190
gi 30409136	KSLKKFHDV L	CEGH	CKNI KLT YFI	MI FASVHFV L	SQLPNFNSI S	177
Consensus	KSL-KFHDV L	VCGD---CEG	C-NI KLT YFI	MI FASVHFV L	SQLPNFNSI S	200
CeresGdna:1440705	GVSLAAAVMS	LSYSTI AWGA	SVHKGVPD V	DYGYKASTTS	GLVFNFFSAL	184
CeresClone:1571117	GVSLAAAVMS	LSYSTI AWGA	SVSKGRVPD V	DYGLRATTTP	GKVFGF GAL	249
Lead:Lead120-CeresClone218626	GVSLAAAVMS	LSYSTI AWGA	SVHKGRMSGV	DYHLRATTTP	GKVFGF GAL	240
gi 30409136	GVSLAAAVMS	LSYSTI AWGA	SVDKGVA D V	DYHLRATTST	GKVFGF SAL	227
Consensus	GVSLAAAVMS	LSYSTI AWGA	SVHKGRV-DV	DY-LRATTT-	GKVFGF -AL	250

Figure 4

CeresGdna:1440705	GDVAFAYAGH	NVVLEIQATI	PSKPKPSKKG	PMWKGVVVAY	I VVALCYFPV	234
CeresClone:1571117	G VAFAYAGH	NVVLEIQATI	PSTPEKPSKK	PMWKGVVVAY	LVVALCYFPV	299
Lead-Lead120-CeresClone218626	GDVAFAYAGH	NVVLEIQATI	PSTPDKPSKK	PMWKGVVVAY	VVVALCYFPV	290
gi 30409136	GDVAFAYAGH	NVVLEIQATI	PSTPEKPSKK	PMWKGVVVAY	I I VALCYFPV	277
Consensus	GDVAFAYAGH	NVVLEIQATI	PSTPEKPSKK	PMWKGVVVAY	I VVALCYFPV	300
CeresGdna:1440705	ALIGYWFNG	KVEDNLI SL	EKPTMLT VAA	NMEVVI HVI G	SYQI YAI PVF	284
CeresClone:1571117	SEVGYWAFGD	SVDGDI LVTL	NRPRWLI ALA	NMMVVI HVI G	SYQI YAMPVF	349
Lead-Lead120-CeresClone218626	ALIGYWAFGN	I VEDNLI LITL	SKPKWLI ALA	NMMVVVHVI G	SYQI YAMPVF	340
gi 30409136	ALVGYWAFGN	HMDDNLI LITL	SRPKWLI ALA	NMMVVI HVI G	SYQI YAMPVF	327
Consensus	AL -GYWAFGN	-V -DNI LITL	S -PKWLI ALA	NMMVVI HVI G	SYQI YAMPVF	350
CeresGdna:1440705	DMLETLVKK	LHFRPSRKLR	FTRNYVAF	TMFVGI CFPF	FGGLLGFFGG	334
CeresClone:1571117	DMIEVLVKK	LRFPPLTLR	LIARTVYVAF	TMFIAI TFPF	F DGLLSFFGG	399
Lead-Lead120-CeresClone218626	DMIEVLVKK	LRFPPLTLR	LIARTLYVAF	TMFIAI TFPF	FGGLLGFFGG	390
gi 30409136	DMIEVLVKK	LRFPPLTLR	LIARTLYVAF	TMFIAI TFPF	FGGLLGFFGG	377
Consensus	DMIEVLVKK	LRFPPLTLR	LIARTIYVAF	TMFIAI TFPF	FGGLLGFFGG	400
CeresGdna:1440705	FAFAPTTYFL	PCIMWLAIYK	PKRFSLSWLT	NWCLDLRTNL	VSDSLFPMKR	384
CeresClone:1571117	FAFAPTTYFL	PCIMWLAIYK	PKRFSLSWFT	NW	---	431
Lead-Lead120-CeresClone218626	FAFAPTTYFL	PCVMWLAIYK	PKRFSLSWLT	NW	---	422
gi 30409136	FAFAPTTYFL	PCIMWLAIYK	PRRFSLSWFT	NW	---	409
Consensus	FAFAPTTYFL	PCIMWLAIYK	PKRFSLSW-T	NW	---	450
CeresGdna:1440705	HGRNSGVTLI	SSTFCITGLGG	HDSSLSFSQL	I SRFCGRAYS	CNVD LVRFP	434
CeresClone:1571117	---	CIILG	---	---	VLLMVLAP	445
Lead-Lead120-CeresClone218626	---	MCIILG	---	---	VLLMI LSP	436
gi 30409136	---	CIILG	---	---	VMLMI LSP	423
Consensus	---	CIILG	---	---	VLLM-LSP	500

Figure 4 (continued)

CeresGdna:1440705	VVWYREFSH	QKLLKS	461
CeresClone:1571117	IIISAKTYKF	YQ----	463
Lead·Lead120·CeresClone218626	IIIMDAKTYQF	YS----	454
gi 30409136	IIIDAKTYKF	YS----	441
Consensus	I I I DAKTYKF	Y-----	527

Figure 4 (continued)

Lead·CeresClone121021	-----MENGCTI	PGQEIGSGLQ	KPVNPDGF--	-----VTI	28
CeresClone:1121512	SSCFDNGEPL	DSEQKSQKIF	VSDHI NAFQY	ADEKADSFVI	50
CeresGdna:1501628	SSGGDGGLLL	EETMEAQKLT	VLDRI DGFQY	TKEKSDSFI I	42
Consensus	SS--DNG--L	-----QKL-	V-D-I DGFQY	--EK-DSF-I	50
Lead·CeresClone121021	DVESFSSVIH	KDFSSSSPRI	TLQRNVSRKG	SPRSNNER--	75
CeresClone:1121512	DMDAFSSGHN	KDAINANSRI	TLQRSLSRKG	SQRLGDMKLN	100
CeresGdna:1501628	DMESFSHGIN	KDJ-NTNQRI	T--RNLSRKG	SPRCGGGG--	77
Consensus	DMESFSSGIN	KD--N-N-RI	TLQRNLSRKG	SPR-G-----	100
Lead·CeresClone121021	DKEISFPQSP	LRCSSSTPEKA	S--TVGPTIEH	AGTATTAATA	123
CeresClone:1121512	VPAICSPKXT	LVGPFITPEKP	AGMAVGPIMGH	SMNPHVH---	138
CeresGdna:1501628	-----	GPSTPEKA	AMVTVGTPDH	SSSPQVH---	106
Consensus	---I---P---	L-GPSTPEKA	A--TVGP--H	S--P-VH---	150
Lead·CeresClone121021	VTTAATAAGN	MITDQNRERR	FGFSRKSSFK	RSHTSWMLDP	173
CeresClone:1121512	LTADNIP T--	-----ESK	CSITRRNSFR	RPS-SWAJ DP	178
CeresGdna:1501628	LTGSLNG--	-----TPEGR	C--I RRRNSFK	RASPSWVLDP	147
Consensus	-TT--I-----	-----E-R	C---RRNSFK	R-S-SW-LDP	200
Lead·CeresClone121021	SSMGSI LLI	FTLSISKSNP	GDMPLD		199
CeresClone:1121512	SSMGTM LLI Y	FTLTI SKQSA	EEYGG--		203
CeresGdna:1501628	SSMGTM LLI Y	LTLSIGK LKT	D-----		168
Consensus	SSMGTM LLI Y	FTLSISK---	-----		226

Figure 5

CeresClone:754061  
 CeresClone:282892  
 gil50925813  
 Lead-CeresClone158765  
 gil5669656  
 CeresClone:537752

MMMTQV-ANH  
 MMMAEVAANH  
 MMTTAV-ANH  
 ---MVFSGDA  
 MFSRREGIDV  
 ---

SKRNHITDGYF  
 SKRSHNDGYF  
 SKRNHNEGYF  
 GMMMEENK  
 ---MPHK  
 RKMMEHK  
 ---

SGKQQQQQQ  
 SAK  
 T  
 ---

VGAITATSSG  
 ---AAAAAAS  
 ---GKAATSS  
 ---RNVCSLGE  
 ---RSPI SLEH  
 ---RRPCSV DQ  
 ---

SEEFGL--RMG  
 PEEELGSMMS  
 SEEFGL--SMT  
 ---S  
 SSSLT--SLT  
 SSYT--SIA  
 ---

47  
41  
36  
24  
20  
32  
50

Consensus  
 ---R--AS--SM--

CeresClone:754061  
 CeresClone:282892  
 gil50925813  
 Lead-CeresClone158765  
 gil5669656  
 CeresClone:537752

SKKPRSA SPR  
 SKKPRNSNSP  
 SKKPRNTSPR  
 ---LRFNLSKE  
 PKRLKAD  
 SKRQKAD  
 ---

RT-AVSPKE  
 D--AVSPKE  
 ---LRFNLSKE  
 ---MLTSSKE  
 ---LSLSTKE  
 ---

KKDKVGERVA  
 KKDRIGERVA  
 KKDKIGERVA  
 RKDKVGERIS  
 KKEKFGGERIV  
 RKEKIGERIV  
 ---

ALQQLVSPFG  
 ALQQLVSPFG  
 ALQQLVSPFG  
 ALQQLVSPYG  
 ALQQLVSPYG  
 ALQQLVSPYG  
 ---

KTDTASVLQE  
 KTDTASVLQE  
 KTDTASVLQE  
 KTDTASVLQE  
 KTDTASVLQE  
 KTDTASVLQE  
 ---

97  
90  
84  
68  
64  
76  
100

Consensus  
 ---LPI S-KE

CeresClone:754061  
 CeresClone:282892  
 gil50925813  
 Lead-CeresClone158765  
 gil5669656  
 CeresClone:537752

ASGYI KFLHQ  
 ASCYI RFLHQ  
 ASGYI KFLHQ  
 AMHYI EFLHE  
 AMGYI KFLHE  
 AMEYI GFLHK  
 ---

QLEVLSSPYM  
 QLQVLSSPYM  
 QLEVLSSPYM  
 QVKVLSAPYL  
 QVKVLSAPYL  
 QVKLLSAPYL  
 ---

RPPA-PCAEP  
 RAPPAA-CAAP  
 RAPPVP-CAAP  
 QTTIPD--ATQ  
 GTMPM--SKT  
 ESSPA--AKM  
 ---

EDPEHYSLRN  
 EDTEHYSLRS  
 EDPEHYSLRN  
 EELQYSLRN  
 QESQPNLRS  
 QGVPEPCSLRS  
 ---

RGLCLVPVEQ  
 RGLCLVPVDQ  
 RGLCLVPVDQ  
 RGLCLVPMEIN  
 QGLCLVPVSY  
 RGLCLVPVSV  
 ---

147  
140  
134  
116  
112  
124  
150

Consensus  
 A-GYI KFLHQ Q-KVLS-PY- R-PPA-GAKP ED-EHYSLR- RGLCLVPV-Q

CeresClone:754061  
 CeresClone:282892  
 gil50925813  
 Lead-CeresClone158765  
 gil5669656  
 CeresClone:537752

TLQLTQSNGA  
 TLQLTQSNGA  
 TLQLTQSNGA  
 TVGVAQSNGA  
 TVGVAATSNGA  
 TI GVAESNGA  
 ---

DLWAPANTSR  
 DLWAPANTTR  
 DLWAPANTTR  
 DI WAPVKTPL  
 DI WAPVKTSL  
 DI WAPVKTIT  
 ---

RT  
 RR  
 RR  
 SPAFSVTSQS  
 SSSPE-NDV  
 SPKFE-KDVS  
 ---

169  
162  
156  
149  
140  
156  
183

Consensus  
 TL---QSNGA D-WAP---TTR

Figure 6

Lead·CeresClone16403	MAVSSLSI RC	GGFSPTI SHK	TE LCPNPS	KACCLLSSGG	KADSESTYQ	50
CeresClone:611156	MVSSCSL --	SW SPCLSHK	LNLPHITN-CL	PRNIATSSSN	TVFCELDITP	47
CeresGdna:1464944	MAISSLSL --	SWASTLSQK	LSVPGSNEL	PRVAFSGNN	SVTCTAEATF	48
Consensus	MAVSSLSL --	SW-SPTLSHK	L--P--N--L	PR--A--SS--N	-V-C---IT-	50
Lead·CeresClone16403	KGSGNWKRR	QALVGVGILV	ATSPATLL	AEEIPKYSIP	FVDR <del>E</del> DGYSY	100
CeresClone:611156	SGESHCRRR	PLLLGI GALT	ANLQPTNLVF	AQEKPRYRA	FVDYEDGYSY	96
CeresGdna:1464944	NEESNCRRR	LLLLGVGALT	TSLVPANFLF	AEEIPKNYTS	FVDFEDGYSY	97
Consensus	-GESNCRRR	-LLLGVGALT	A-L-PANLLF	AEEIPK-Y--	FVD-EDGYSY	100
Lead·CeresClone16403	YYP <del>S</del> DWREFD	FRAHDSAFKD	RYLQ <del>L</del> QNVRV	RFI <del>P</del> TEKNDI	HEVGPMEEVV	150
CeresClone:611156	VY <del>P</del> DWKEFD	FRAHDSAFKD	RYLQ <del>L</del> QNVRV	RFI <del>P</del> TEKKDI	RDLGPMEEVI	146
CeresGdna:1464944	YYP <del>S</del> DWDFD	FRGHDSAFKD	R <del>F</del> K <del>L</del> Q <del>L</del> QNVRV	RFI <del>P</del> TEKKDI	HELGPMEE--	145
Consensus	YYP <del>S</del> DW-EFD	FRAHDSAFKD	RYLQ <del>L</del> QNVRV	RFI <del>P</del> TEKKDI	HELGPMEEV-	150
Lead·CeresClone16403	YDLVKHKFAA	PNQVATI YDM	KERVE <del>D</del> GK <del>N</del> Y	YTFEYGLRTP	IYATTSFATV	200
CeresClone:611156	YDLVKHRYAA	PNQRPTI NDM	QEKTI DGKHY	YTFEYLTSP	NYSSASFATI	196
CeresGdna:1464944	YDSHM <del>Q</del> QEI M	NV <del>K</del> LSNFLE-	NQKTVEGK <del>N</del> Y	YTFEYELTSP	NYSSVSFATI	194
Consensus	YDLVKH--AA	PNQ--TI -DM	--EKT-DGK <del>N</del> Y	YTFEY-LTSP	NYSS-SFATI	200
Lead·CeresClone16403	AVGN <del>N</del> RYTYTL	I VGANERRWR	KV <del>K</del> QLQVVA	DSL <del>K</del> ILQJ	238	
CeresClone:611156	AI GNGRYTYTL	I VGANERRWK	RFRDQLKVVA	DSFRLLDI	234	
CeresGdna:1464944	VI ANGRFYTL	I VGANERRWR	RYRSQLKVVA	DSFKVLDI	232	
Consensus	AI GNGRYTYL	I VGANERRWR	R-R-QLKVVA	DSFK-LDI	238	

Figure 7

Lead:CeresClone28635	50	YSMLHKVSR				
gij38426486	49	YSMLHKVSR	PPEPHWFC			
CeresGdna:1514021	29	YSMLHRVSR	PSQPHWAFS			
gij55710094	50	YSMLHKVSR	PPEPHWFC			
gij75859951	50	YSMLHKVSR	PPEPHWFC			
gij41224629	50	YSMLHKVSR	PPEPHWFC			
gij27475614	50	YSMLHKVSR	PPEPHWFC			
gij28208268	50	YSMLHKVSR	PPEPHWFC			
gij1449163	50	YSMLHKVSR	PPEPHWFC			
gij2463569	50	YSMLHKVSR	PPEPHWFC			
Consensus	50	YSMLHKVSR	I PPEPHWFC			
Lead:CeresClone28635	100	VPI LI AFHRH	DDTSI PTDEK	LVL RALDTVE	ELR NAVCVFY	FSLVI QQLNT
gij38426486	99	VPI LI AFHKH	DDTSI AADTK	LVL RALDTVE	QLR DAVCI FY	FALVI QQLNP
CeresGdna:1514021	79	VPI LI AFHRH	DDTSI PTDVK	LVL RALDTVE	ELR NAVCI FY	FALVI QQLGI
gij55710094	100	VPI LKAFHRH	DDTSI STEVK	LVL RALDTVE	QLR DAVCI FY	FGLVI QQLGP
gij75859951	100	VPI LMAFHRH	DDTSI PTEVK	LVL RALDTVE	QLR DAVCI FY	FGLVI QQLGP
gij41224629	100	VPI LMAFHRH	DDTSI PTEVK	LVL RALDTVE	QLR DAVCI FY	FGLVI QQLGP
gij27475614	100	VPI LI DFHRH	DDTSI ETDVK	LVL RALDTVE	ELR DAVCI FY	FGLVI QQLGP
gij28208268	100	VPI LKAFHRH	DDTSI ATDVK	LVL RALDTVE	DLR NAVCI FY	FALVI QQLDT
gij1449163	100	VPI LI AFHRH	DDTSI ATDVK	LVL RALDTVE	QLR NAVCI FY	FALVI QQLDP
gij2463569	100	VPI LI AFHRH	DDTSI ETDVK	LVL RALDTVE	ELR NAVCI FY	FALVI QQLGI
Consensus	100	VPI LI AFHRH	DDTSI -TDVK	LVL RALDTVE	QLR -AVCI FY	FALVI QQLGP
Lead:CeresClone28635	150	EET TRRMGAG	ELKRGYQEA	QFHHVSAFLL	GTKEYKI LMD	YDIDWHYSC
gij38426486	149	EDI TKRMGAG	ELKRGYQEA	QFHHVSTAFLL	GTKEYKVLMD	YNRDWHFAC
CeresGdna:1514021	129	EDI TKRMGAG	ELKRGYQEA	QFHDVSTAFLL	GTNDYRVLMD	YDRNWHFSC
gij55710094	150	EDI TKRMGAG	ELSGYQEA	EFHHVSNFLL	GTKEYKI LMD	YDNNWHFSC
gij75859951	150	EDI TKRMGAG	ELSGYQEA	EFHHVSNFLL	GTKEYKVLMD	YDKDWHFSC
gij41224629	150	EDI TKRMGAG	ELSGYQEA	EFHHVSNFLL	GTKEYKVLMD	YDKDWHFSC
gij27475614	150	EDI TKRMGAG	ELGKNTQDA	QFHHVSKAFL	GTKEYKVLMD	YDNDWHFSC
gij28208268	150	EDI TKRMGAG	ELAKNYQEA	QFHDVSTAFLL	GTKEYKVLMD	YDRDWHFSC
gij1449163	150	ENI TKRMGAG	ELKNYQEA	QFHHVSTAFLL	GTKEYKVLMD	YDRDWHFSC
gij2463569	150	EDI TKRMGAG	ELKNTQDA	QFHHVSTAFLL	GTKEYKVLMD	YDRDWHFSC
Consensus	150	EDI TKRMGAG	ELGKGYQEA	QFHHVSTAFLL	GTKEYKVLMD	I YDRDWHFSC

Figure 8



Lead:CeresClone28635	MAKFI COEVE	TVDDYDEYCH	YVAGLVGLGL	SKLFLAAGSE	VITPDMEATS	200
gi 38426486	MAKFI CKEVE	TVDDYDEYCH	YVAGLVGLGL	SKLFHSSGTE	ILFSD--SIS	197
CeresGdna:1514021	MAKFI CKEVE	TI DDYDEYCH	YVAGLVGLGL	SKLFHASGE	DLASD--SIS	177
gi 55710094	MAKFI CKEVE	TI DDYDEYCH	YVAGLVGLGL	SKLFHASGAE	DLASD--SLS	198
gi 75859951	MAKFI CKEVE	TI DDYDEYCH	YVAGLVGLGL	SKLFHASGAE	DLATD--SLS	198
gi 41224629	MAKFI CKEVE	TI DDYDEYCH	YVAGLVGLGL	SKLFHASGAE	DLATD--SLS	198
gi 27475614	MAKFI CKEVE	TI DDYDEYCH	YVAGLVGLGL	SKLFHASGAE	DLATD--SLS	198
gi 28208268	MAKFI CKEVE	TI DDYDEYCH	YVAGLVGLGL	SKLFHASGAE	DLATD--SLS	198
gi 1449163	MAKFI CKEVE	TI DDYDEYCH	YVAGLVGLGL	SKLFHASGAE	NLAAD--SLS	198
gi 2463569	MAKFI CKEVE	TI DDYDEYCH	YVAGLVGLGL	SKLFHASGAE	DLAPD--HLS	198
Consensus	MAKFI CKEVE	TI DDYDEYCH	YVAGLVGLGL	SKLFHASG-E	DLA-D--SLS	200
Lead:CeresClone28635	NSMGLFLQKT	NIIRDYLEDI	NEIPKSRMFW	PREIWKYAD	KLEDLKYEEN	250
gi 38426486	NSMGLFLQKT	NIIRDYLEDI	NEIPKSRMFW	PREIWSKYVN	KLEDLKYEEN	247
CeresGdna:1514021	NSMGLFLQKT	NIIRDYLEDI	NEIPKSRMFW	PREIWSKYVN	KLEDLKYEEN	227
gi 55710094	NSMGLFLQKT	NIIRDYLEDI	NEIPKSRMFW	PRKIWNKYVN	KLEDLKYEEN	248
gi 75859951	NSMGLFLQKT	NIIRDYLEDI	NEIPKSRMFW	PROIWSKYVD	KLEDLKYEEN	248
gi 41224629	NSMGLFLQKT	NIIRDYLEDI	NEIPKSRMFW	PROIWSKYVD	KLEDLKYEEN	248
gi 27475614	NSMGLFLQKT	NIIRDYLEDI	NEIPKSRMFW	PROIWSKYVD	KLEDLKYEEN	248
gi 28208268	NSMGLFLQKT	NIIRDYLEDI	NEIPKSRMFW	PROIWSKYVS	KLEDLKYEEN	248
gi 1449163	NSMGLFLQKT	NIIRDYLEDI	NEIPKSRMFW	PROIWSKYVN	KLEDLKYEEN	248
gi 2463569	NSMGLFLQKT	NIIRDYLEDI	NEIPKSRMFW	PROIWSKYVN	KLEDLKYEEN	248
Consensus	NSMGLFLQKT	NIIRDYLEDI	NEIPKSRMFW	PROIWSKYVN	KLEDLKYEEN	250
Lead:CeresClone28635	SVKAVQCLND	MVTNALLHAE	DCLKYMSALR	DPSIFRFCAI	PQIMAI GTLA	300
gi 38426486	SVKAVQCLND	MVTNALLHAE	DCLKYMSALR	DPAI FRFCAI	PQIMAI GTLA	297
CeresGdna:1514021	SVKAVQCLND	MVTNALLHAE	DCLKYMSALR	EPAI FRFCAI	PQVMAI GTLA	277
gi 55710094	SVKAVQCLND	MVTNALLHAE	DCLKYMSALR	DPAI FRFCAI	PQIMAI GTLA	298
gi 75859951	SVKAVQCLND	MVTNALLHAE	DCLKYMSALR	DPAI FRFCAI	PQIMAI GTLA	298
gi 41224629	SVKAVQCLND	MVTNALLHAE	DCLKYMSALR	DPAI FRFCAI	PQIMAI GTLA	298
gi 27475614	SVKAVQCLND	MVTNALLHAE	DCLKYMSALR	DPAI FRFCAI	PQIMAI GTLA	298
gi 28208268	SVKAVQCLND	MVTNALLHAE	DCLKYMSALR	DPAI FRFCAI	PQIMAI GTLA	298
gi 1449163	SVKAVQCLND	MVTNALLHAE	DCLKYMSALR	DPAI FRFCAI	PQIMAI GTLA	298
gi 2463569	SVKAVQCLND	MVTNALLHAE	DCLKYMSALR	DPAI FRFCAI	PQIMAI GTLA	298
Consensus	SVKAVQCLND	MVTNALLHAE	DCLKYMSALR	DPAI FRFCAI	PQIMAI GTLA	300

Figure 8 (continued)

Lead-CeresClone28635  
 gi|38426486  
 CeresGdna:1514021  
 gi|55710094  
 gi|75859951  
 gi|41224629  
 gi|27475614  
 gi|28208268  
 gi|1449163  
 gi|2463569

Consensus

L C Y N N I - V F R G V V K M R R G L T A K V I D R T K T M A D V Y G A F Y D F S C M L K T K V D K  
 L C Y N N I E V F R G V V K L R R G L T A K V I D R T K T M A D V Y Q A F S D F S C M L K S K V D M  
 M C Y N N I N V F R G V V K M R R G L T A Q I F H R T K T M A D V Y G A F F D F S C M L K S K V D R  
 L C Y N N L Q V F R G V V K M R R G L T A K V I D R T N K M A D V Y G A F Y D F S C M L K T K V D N  
 L C F N N I Q V F R G V V K M R R G L T A K V I D R T K T M A D V Y G A F F D F S C L L K S K V D N  
 L C F N N I Q V F R G V V K M R R G L T A K V I D Q T K T M A D V Y G A F F D F S C L L K S K V D N  
 M C Y N N I Q V F R G V V K M R R G L T A K V I D R T K T M A D V Y G A F F D F A S V L E S K V D K  
 L C Y N N V E V F R G V V K M R R G L T A K V I D R T K T I A S M L E S K V D K  
 L C Y N N I E V F R G V V K M R R G L T A K V I D R T K T M A D V Y G A F F D F S C M L K S K V D K  
 L C Y N N I E V F R G V V K M R R G L T A K V I D R T K T M A D V Y G A F F D F A S M L E P K V D K  
 L C Y N N I - V F R G V V K M R R G L T A K V I D R T K T M A D V Y G A F F D F S C M L K S K V D K  
 Consensus 350

Lead-CeresClone28635  
 gi|38426486  
 CeresGdna:1514021  
 gi|55710094  
 gi|75859951  
 gi|41224629  
 gi|27475614  
 gi|28208268  
 gi|1449163  
 gi|2463569

Consensus

N D P N A S K T L N R L E A V Q K L C R D A G V L Q - N R K S Y V N D K - - - - - G Q P N S V F  
 H D P N A Q T I T R L E A A Q K I C K D S G T L S - N R K S Y I V K R E - - - - - S S Y S A A L  
 N D P N A T K T L S S L E A V Q K T C R E S G A L N - K R M V D F L N S F V C V W M M A G Y T S A C  
 N D P N A T K T L S R L E A I Q K K C K E S G V I T P N R K S Y V L E N D - - - - - S G Y N L V I  
 N D P N A T K T L S R L E A I Q K T C K E S G T L S - K R K S Y I I E S E - - - - - S G H N S A L  
 N D P N A T K T L S R L E A I Q K T C K E S G T L S - K R K S Y I I E S E - - - - - S G H N S A L  
 N D P N A T K T L S S S R L E A I Q K T C R E S G L L T - K R K S Y V L R N E - - - - - S G Y G S T I M  
 N D P N A T K T L S R L V A I Q K T C R E S G L L N - K R K S Y I L R K E - - - - - N G Y G S T I M  
 N D P N A T K T L S R L E A I Q K T C R E S G L L S - K R K P Y I L R N E - - - - - S T N S S T I M  
 N D P N A T K T L S R L E A I Q K T C R E S G L L S - K R K S Y I V N D E - - - - - S G Y G S T I M  
 N D P N A T K T L S R L E A I Q K T C R E S G L L S - K R K S Y I L - - - - - S G Y - S - L  
 Consensus 400

Lead-CeresClone28635  
 gi|38426486  
 CeresGdna:1514021  
 gi|55710094  
 gi|75859951  
 gi|41224629  
 gi|27475614  
 gi|28208268  
 gi|1449163  
 gi|2463569

Consensus

I I M V V I L L A I V F A Y L R A N - - - - - 410  
 L A L L F T I L A I L Y A Y L S A N R P I K I - - - - - 413  
 S I P W K L L L L - - - - - 385  
 I A I L F I L L A L V Y A Y L S S N L S N N R - - - - - 415  
 I A I F I I L A I L Y A Y L S S N L L P N K O - - - - - 415  
 I A I F I I L A I L Y A Y L S S N L L L N K O - - - - - 415  
 I L L V I L F S I I F A Y L S A N R H N N - - - - - 413  
 I I L V L L F S I M F A Y S A I T R H S N - - - - - 413  
 M L I L V I L L S I I F A Y L S A K R Q D N - - - - - 413  
 I V I L V I M V S I I F A Y L S A N H H N S - - - - - 413  
 I I I L V I I L A I - Y A Y L S A N - - - - - 424  
 Consensus

Figure 8 (continued)

g 4100646.T	---	MAA	KRI	KELQDL	QRPP	SCSI	GPV	GDLFHW	33
g 52851174.T	---	MAS	RRI	LKELREL	QRPP	SCSA	GPVA	QDMFHW	33
g 54402104.T	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPV	DDMFHW	33
g 30025160.T	---	MAS	KRI	LKEFKDL	QKDP	PPVSCSA	GPV	GEDMFHW	33
Clone:617835.T	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPV	GEDMFHW	33
Gdna:1483290.T	MR	RRINPVLA	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	50
g 22597164.T	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	33
g 441457.T	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	33
Clone:1380019.T	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	33
Lead:Clone35698	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	33
g 5762457.T	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	33
g 50906823.T	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	33
g 77416935.T	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	33
g 40287554.T	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	33
g 28569265.T	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	33
g 28569267.T	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	33
Consensus	---	MAS	KRI	LKELKDL	QKDP	PPVSCSA	GPVA	EEDMFHW	50

g 4100646.T	QATI	MGPVDS	HFPP	DDY	PPFK	PPKV	IF	FKTKV	YHPNV	NSNGS	83
g 52851174.T	QATI	I	GP	NDS	HFPP	DDY	PPFK	PPKV	AF	RTKV	83
g 54402104.T	QATI	MGP	ADS	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
g 30025160.T	QATI	MGP	HDS	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
Clone:617835.T	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
Gdna:1483290.T	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	100
g 22597164.T	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
g 441457.T	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
Clone:1380019.T	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
Lead:Clone35698	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
g 5762457.T	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
g 50906823.T	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
g 77416935.T	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
g 40287554.T	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
g 28569265.T	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
g 28569267.T	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	83
Consensus	QATI	MGP	S	HFPP	DDY	PPFK	PPKV	SF	KT	KV	100

Figure 9

gi 4100646.T	CLDI LKEQW	SPALTVSKV	---	102
gi 52851174.T	CLDI LKDQW	SPALTI SKV	---	102
gi 54402104.T	CLDI LKEQW	SPALTVSKV	---	102
gi 30025160.T	CLDI LKEQW	SPALTI SKV	---	102
Clone:617835.T	CLDI LKDQW	SPALTI SKV	---	102
Gdno:1483290.T	CLDI LKEQW	SPALTI SKV	---	119
gi 22597164.T	CLDI LKEQW	SPALTI SKV	---	102
gi 441457.T	CLDI LKEQW	SPALTI SKV	---	102
Clone:1380019.T	CLDI LKEQW	SPALTI SKV	---	102
Lead.Clone35698	CLDI LKEQW	SPALTI SKV	FRYVHC	109
gi 5762457.T	CLDI LKEQW	SPALTI SKV	---	102
gi 50906823.T	CLDI LKEQW	SPALTI SKV	---	102
gi 77416935.T	CLDI LKEQW	SPALTI SKV	---	102
gi 40287554.T	CLDI LKEQW	SPALTI SKV	---	102
gi 28569265.T	CLDI LKEQW	SPALTI SKV	---	102
gi 28569267.T	CLDI LKEQW	SPALTI SKV	---	102
Consensus	I CLDI LKEQW	SPALTI SKV	---	126

Figure 9 (continued)

Lead:CeresClone36412	MAVDAHHLFL	SPPQLFSNRE	LT	MNNNTMEP	T	SGGFCN	N	N	Q	T	C	G	V	S	P	F	S	50	
CeresGdna:1467033	-----	-----	---	MMN	---	---	MYN	T	Q	---	M	C	Y	---	---	---	---	9	
CeresClone:1641329	-----	-----	---	MMN	---	---	MYNS	T	Q	---	M	C	Y	S	S	F	---	17	
Consensus	-----	-----	---	MN	---	---	MYNQ	T	Q	---	M	C	Y	S	S	F	---	50	
Lead:CeresClone36412	VPNHTSTTTT	ATPPLHMYG	G	SDTIPTIT	AGY	Y	AD	C	A	T	N	L	D	C	E	F	F	97	
CeresGdna:1467033	VPLSGTITTT	TADFLPMYS	S	VITDSISHK	TPI	---	K	S	D	S	G	L	T	Y	N	---	---	54	
CeresClone:1641329	LPLSGAVTEL	ALPS--SVFN	H	S	L	A	N	A	V	K	S	E	N	N	N	V	S	64	
Consensus	VPLSG-TTTT	A-P-LL-MY-	---	D-I	---	---	G	---	Y	-D	---	N	N	---	---	---	100		
Lead:CeresClone36412	TRKRSRDSR	SNYHLLQN	P	R	S	S	C	V	N	A	T	T	T	T	T	P	F	S	147
CeresGdna:1467033	QRKRPRDSN	PLLSYPTL	---	---	---	---	---	---	---	---	Q	S	N	K	T	C	A	P	91
CeresClone:1641329	SRKRSRDN-N	NNYGY	---	---	---	---	---	---	---	---	N	N	N	N	D	S	F	S	95
Consensus	-RKRSDS-N	-NY-Y--L--	---	---	---	---	---	---	---	---	N	T	---	P	F	S	---	150	
Lead:CeresClone36412	HNNQQHEID	RFVSLHLYQM	E	R	V	K	Y	E	I	E	E	E	K	R	Q	R	R	I	197
CeresGdna:1467033	QIQQQQLDI	CLVSOH--M	E	K	V	R	M	E	I	E	E	E	K	R	Q	R	R	I	138
CeresClone:1641329	QIQQQQLDI	HLIMQR--M	E	K	V	R	M	E	I	E	E	E	K	R	Q	R	R	I	142
Consensus	QIQQQQLDI	-LVSQH--M	E	K	V	R	M	E	I	E	E	E	K	R	Q	R	R	I	200
Lead:CeresClone36412	RVKEEERERI	GKVNHALLEER	V	K	S	L	S	E	N	Q	W	R	D	L	A	Q	T	N	247
CeresGdna:1467033	RAKEEIEKI	GKLNWALEEK	V	K	S	L	C	V	E	N	Q	W	R	D	L	A	Q	S	188
CeresClone:1641329	KIKKEEIEKI	GKLNWALEEK	V	K	H	L	C	M	E	N	Q	W	R	N	I	A	E	T	192
Consensus	R-KEEIEKI	GKLNWALEEK	V	K	S	L	C	-E	N	Q	W	R	D	L	A	Q	T	N	250
Lead:CeresClone36412	HVLAQV----	KDMSRCAGLE	K	N	M	N	E	E	D	D	A	E	S	C	C	C	S	C	289
CeresGdna:1467033	QVLAQV--VK	EDRTLGAGLD	D	Q	T	A	A	L	D	D	A	Q	S	C	C	C	S	N	236
CeresClone:1641329	QVLAQRGGMA	AEEIDVGGGAT	V	C	G	G	A	E	M	D	D	A	E	S	C	C	C	S	239
Consensus	QVLAQ-----	-D---GAGL-	---	A	E	-D	D	A	E	S	C	C	C	S	---	C	D	300	

Figure 10

Lead-CeresClone36412  
 CeresGdna:1467033  
 CeresClone:1641329  
 Consensus

```

RRVGLER--E A QDKAERRR R-----MCR NC G E E E S C V L L P C R H L C L C 331
ERCILAN--G A QDNNGAGPR GTGTSSWLCR NC N K A E S C V L L P C R H L C L C 284
GWRITLACCA G V K D K E G G N G R-----LCR NC R K E E S C V L L P C R H L C L C 283
-R-TLA---G A QDK-G-G-R R-----LCR NC-KEESC V L L P C R H L C L C 350
  
```

Lead-CeresClone36412  
 CeresGdna:1467033  
 CeresClone:1641329  
 Consensus

```

G V C G S S V H T C P I C I S P K N A S V H V N M S S 358
I V C G S S L H T C P I C K A T K N A S V H V N L S - 310
I V C G S S L H T C P I C K S Y K T A S V H V N M S - 309
T V C G S S L H T C P I C K S - K N A S V H V N M S - 377
  
```

Figure 10 (continued)



Lead·CeresClone5426	MATVPGQLIW	EIVKRNNCFI	VKQFGRGNAK	VQFSKESNNL	VNI NSYKHSG	50
CeresClone:1123542	MATVPGQLIW	EIVK <del>T</del> NNCFI	VKQFGRGNAK	VQFSKETNNL	CNL NSYKHSG	50
CeresGdna:1499194	MATVPGQLIW	EIVKKNCFI	VKQFGRGTAS	LRFSKESNNL	YNL NSYKHSG	50
CeresClone:557065	MATVPGQLIW	EVVKKN <del>S</del> FI	VKEFGRT <del>QS</del>	VQFSRE <del>P</del> NNL	YNLNTFKYSG	50
Consensus	MATVPGQLIW	EIVKKNCFI	VKQFGRG-A-	VQFSKESNNL	YNLNSYKHSG	50
Lead·CeresClone5426	LANKKTVTIQ	AAGKDQGVVL	GTTKT <del>K</del> RQNK	PKLSVNKSL	KKEFSRMSKV	100
CeresClone:1123542	LANKKTVTIQ	AADKEQGVVL	ATTKT <del>K</del> KQNK	PK <del>M</del> SVNKSIL	KKEFSRMSKA	100
CeresGdna:1499194	LANKKTVTIQ	A <del>G</del> DKDQ <del>A</del> VVL	ATSKTKKQNK	PAAL <del>F</del> HKSVM	KKEFRRMAKA	100
CeresClone:557065	LANKKTVTVQ	PIAGKDQ <del>S</del> VLL	ATTKT <del>K</del> KQNK	PAALLHKSVM	KKEFRRMAKA	100
Consensus	LANKKTVTIQ	AA-KDQ-VVL	ATTKT <del>K</del> KQNK	P---V-KS---	KKEF-RM-KA	100
Lead·CeresClone5426	VANQVVDNYY	RPDLKKAALA	RLSAISK <del>G</del> LR	VAKSG <del>P</del> KRRN	RQA-----	143
CeresClone:1123542	VANQVVDNYY	RPDLK <del>K</del> FALA	RLS <del>M</del> SKSLR	VAKSG-----	-----	135
CeresGdna:1499194	VE <del>N</del> QV <del>E</del> GNYY	RPDLKKAALA	RLSAVHRSLK	VAKSGVKKRN	RQ <del>G</del> LRNFKHG	150
CeresClone:557065	VQ <del>N</del> QV <del>A</del> DNYY	RPDLKKAALA	RLSAVNRSLR	VAKSGVKKRN	RQAVK-----	145
Consensus	V-NQVVDNYY	RPDLKKAALA	RLSA-S-SLR	VAKSGVKKRN	RQA-----	150
Lead·CeresClone5426	-----	-----	-----	-----	-----	143
CeresClone:1123542	-----	-----	-----	-----	-----	135
CeresGdna:1499194	HRNELI GLIF	KACYKN <del>K</del>	-----	-----	-----	167
CeresClone:557065	-----	---VPSR <del>K</del>	-----	-----	-----	150
Consensus	-----	-----K	-----	-----	-----	167

Figure 12



gi|21319                  MAGRNFKNG          ---ENNKAGSS          LHEKNTTP          P          EEAITLKKLG          39  
gi|5823000                MEGRNVKNGD        TNNNKVGSS          LSI EKPI SP        P          -EASSLKKLIA        40  
gi|439294                ---MENG            -TKREGLGKL        TVSSSLQVEQ        P          LAPSKLWKI          34  
gi|575351                ---MENC            -TKKLN            ---SSLAVEQ        P          LPKSKLWKI          28  
Lead-CeresClone7894      MVSHPMKA          -ANGASA           LETQICELDQ        P          ---ERLRKI          33  
gi|18091781              ---MEKC            -SNGSAA           LEMQISELDQ        P          ---AALRKI          28  
gi|415988                ---MENS            -NRGVCA           -KTITTP           P          EEAAPRNF          30  
gi|6120115                ---MENG            ---DRTAP           -AFLOQASP        P          -EAAPVRI          29  
gi|17447420              ---MEVG            ---KSSVF           -NEAKSIALP        P          AQAAPVKNII        24  
gi|68161544              ---MELA            -SANKDM           -AIQDHQSS        P          ---PTPIWKITV       28  
gi|51863031              ---MEVE            -SSTKQN           -RA                P          AQRSTLRIV        22  
gi|33620334              ---MEPF            -SSTKQI           -HNNNITLTK        P          LEPSPLRKI        38  
gi|5230818                ---MEPL            -SSTKQI           -NNNNNLAK        P          LEPSPLRKM        39  
gi|1935019                ---MEPL            -SSTKQI           -NNNNNLAK        P          LEPSPLRKM        38  
gi|7649151                ---METQ            -LATKNP           -SQQPPPP         P          PPSNSLKKIV        33  
gi|468562                ---MQSST           SKENKQP           ---PSSQPH         P          PNSSPLRKVV        39  
CeresGdna:1479767        -----ME-----S--K-----SN-----S-LRKI      50

gi|21319                  LVASVAAGVQ        FGWALQLSLL        TPYVQLLGI P        HTWAAFIWLC        GPI SGMI VQP        89  
gi|5823000                LVASIAAGVQ        FGWALQLSLL        TPYVQLLGI P        HTWAPYI WLC        GPI SGMI VQP        90  
gi|439294                VVASIAAGVQ        FGWALQLSLL        TPYVQLLGI P        HKFASFI WLC        GPI SGMI VQP        84  
gi|575351                MVASIAAGVQ        FGWALQLSLL        TPYVQLLGI P        HKFASFI WLC        GPI SGMI VQP        78  
Lead-CeresClone7894      SVSSIAGVQ         FGWALQLSLL        TPYVQLLGI P        HKWASLI WLC        GPI SGMLVQP        83  
gi|18091781              SVSSIAGVQ         FGWALQLSLL        TPYVQLLGI P        HKWASLI WLC        GPI SGMLVQP        78  
gi|415988                LVAAIAAGVQ        FGWALQLSLL        TPYVQLLGI P        HKWASLI WLC        GPI SGMI VQP        80  
gi|6120115                VVAIAAGVQ         FGWALQLSLL        TPYVQLLGI P        HTWAAFI WLC        GPVSGMLVQP        79  
gi|17447420              VVAIAAGVQ         FGWALQLSLL        TPYVQLLGI P        HVWAAFI WLC        GPI SGLLVQP        74  
gi|68161544              VVASIAGIQ         FGWALQLSLL        TPYVQLLGI P        HKWAAFI WLC        GPI SGMI VQP        78  
gi|51863031                VVASIAAGVQ        FGWALQLSLL        TPYVQLLGI P        HTWAAFIWLC        GPI SGMLVQP        72  
gi|33620334                VVASIAAGVQ        FGWALQLSLL        TPYVQLLGI P        HTWAAFIWLC        GPI SGMLVQP        88  
gi|5230818                VVASIAAGVQ        FGWALQLSLL        TPYVQLLGI P        HTWAAFIWLC        GPI SGMLVQP        89  
gi|1935019                VVASIAAGVQ        FGWALQLSLL        TPYVQLLGI P        HTWAAFIWLC        GPI SGMLVQP        88  
gi|7649151                VVASIAAGVQ        FGWALQLSLL        TPYVQLLGI P        HTWAAFIWLC        GPI SGMLVQP        83  
gi|468562                MVASIAAGIQ        FGWALQLSLL        TPYVQLLGI P        HTWAAFIWLC        GPI SGMLVQP        89  
CeresGdna:1479767        -----VVASI AAGVQ FGWALQLSLL TPYVQLLGI P HTWAAFIWLC GPI SGMLVQP 100  
Consensus

gi|21319  
gi|5823000  
gi|439294  
gi|575351  
Lead-CeresClone7894  
gi|18091781  
gi|415988  
gi|6120115  
gi|17447420  
gi|68161544  
gi|51863031  
gi|33620334  
gi|5230818  
gi|1935019  
gi|7649151  
gi|468562  
CeresGdna:1479767

gi|21319  
gi|5823000  
gi|439294  
gi|575351  
Lead-CeresClone7894  
gi|18091781  
gi|415988  
gi|6120115  
gi|17447420  
gi|68161544  
gi|51863031  
gi|33620334  
gi|5230818  
gi|1935019  
gi|7649151  
gi|468562  
CeresGdna:1479767

Figure 13

gjl21319  
 gjl5823000  
 gjl439294  
 gjl575351  
 Lead-CeresClone7894  
 gjl18091781  
 gjl415988  
 gjl6120115  
 gjl17447420  
 gjl68161544  
 gjl51863031  
 gjl33620334  
 gjl5230818  
 gjl1935019  
 gjl7649151  
 gjl468562  
 CeresGdna:1479767  
 139  
 140  
 134  
 128  
 133  
 128  
 130  
 129  
 124  
 128  
 122  
 138  
 139  
 138  
 133  
 139  
 100

I VGYHSDRCT SRFGRRRPFI AAGA-LVAIA VFLI GYAADL GHASGDSL-K  
 Consensus

gjl21319  
 gjl5823000  
 gjl439294  
 gjl575351  
 Lead-CeresClone7894  
 gjl18091781  
 gjl415988  
 gjl6120115  
 gjl17447420  
 gjl68161544  
 gjl51863031  
 gjl33620334  
 gjl5230818  
 gjl1935019  
 gjl7649151  
 gjl468562  
 CeresGdna:1479767  
 189  
 190  
 184  
 177  
 183  
 178  
 180  
 179  
 174  
 178  
 172  
 188  
 189  
 188  
 183  
 189  
 150

I VGYHSDRCT SRFGRRRPFI AAGA-LVAIA VFLI GYAADL GHASGDSL-K  
 Consensus

Figure 13 (continued)

gi 21319	SFFMALGNIG	GYAAGSYSRL	YTVFPFTKTA	ACDYYCANLK	SCFFISITLL	239
gi 5823000	SFFMALGNIG	GYAAGSYSRL	YTVFPFTKTK	ACDYYCANLK	SCFFISITLL	240
gi 439294	SFFMAVGNIL	GYAAGSYSRL	EKVFPFSKTK	ACDYYCANLK	SCFFIAIFLL	234
gi 575351	SFFMAVGNVL	GYAAGSYSRL	CKI FPFSTP	ACDI YCANLK	SCFFI AVFLL	227
Lead:CeresClone7894	SFFMAVGNVL	GYAAGSYSRNL	YKVVPPFTME	SCDL YCANLK	TCFFLSITLL	233
gi 18091781	SFFMAVGNVL	GYAAGSYSKNL	YKVVPPFTMK	SCDL YCANLK	TCFFLSITLL	228
gi 415988	SFFMAVGNVL	GYAAGSYSRML	YKVFPPFSKTK	ACDI YCANLK	SCFFISITLL	230
gi 6120115	SFFMAVGNVL	GYAAGSYTHL	EKI FPFSTK	ACDYYCANLK	SCFFI SVALL	229
gi 17447420	SFFMAVGNVL	GYAAGSYTRL	YKVFPPFSKTE	ACDYYCANLK	SCFFI SVALL	224
gi 68161544	SFFMAVGNVL	GYAAGSFSKL	HKMFPPARTQ	ACDYYCANLK	SCFFLSIALL	228
gi 51863031	SFFMAVGNVL	GYAAGSYTRF	YKI FPFSTKTK	ACDYYCANLK	SCFFLSIALL	222
gi 33620334	SFFMAVGNVL	GYAAGAYSKL	YHVPFPTKTK	ACDYYCANLK	SCFFLSIALL	238
gi 5230818	SFFMAVGNVL	GYAAGAYSKL	YHVPFPTKTE	ACNYYCANLK	SCFFLSIALL	239
gi 1935019	SFFMAVGNVL	GYAAGAYSKL	YHVPFPTKTK	ACNYYCANLK	SCFFLSIALL	238
gi 7649151	SFFMAVGNIL	GYAAGAYTHL	YKI APFPTKTK	ACDYYCANLK	TCFFISIALL	233
gi 468562	SFFMAVGNVL	GYAAGAYTHL	YKLFPPFTKTK	ACDYYCANLK	SCFFISIMLL	239
CeresGdna:1479767	SFFMAVGNVL	GYAAGSYTHL	YKLFPPFSRTK	ACDYYCANLK	SCFFISIALL	200
Consensus	SFFMAVGNVL	GYAAGSYSRL	YKVFPFTKTK	ACDYYCANLK	SCFFISIALL	250
gi 21319	LTLTILALS	VKERQITIDE	I QEEEELNKR	NN	CARLPFFCQL	284
gi 5823000	LTLTILALS	VREPFITIDE	I QEEEELNKNN	---	CARLPFFCQL	283
gi 439294	LSLTIALTL	VRENELPEKD	EQEI DEKLA	---	KSKVPFFGEI	276
gi 575351	LSLTIALTL	VRENELPEKD	EHEI DEKA	---	KSKVPFFGEI	268
Lead:CeresClone7894	LIVTFVSLCY	VKEKPMTEPP	TADGKA	---	SNVPFFGEI	268
gi 18091781	LIVTFVSLCY	VTEKPMTEPP	TADGKP	---	SNVPFFGEI	263
gi 415988	LTLTILALS	VREKRHVAEE	QVTAAKKG	---	FKI PVFPEL	267
gi 6120115	LCVSAALATI	VRETPPETATA	EAEPAITKK	---	KKI PVFGEI	266
gi 17447420	LITVITIALCI	VRETPYTAPP	EEAGIVKK	---	HTVPVFGEL	261
gi 68161544	LITVITIALCI	VRETPYTAPP	I AVVE	---	AGQPFYSQM	262
gi 51863031	STVTILALTS	VKERALSSQP	KPENAAAGEDE	ER	SAGLPFFGEM	267
gi 33620334	TTLATAALMY	VKEVPLSPEK	AVI DSDDN	---	GGMPCFQOL	275
gi 5230818	TVLATAALY	VKETPLIAEK	AVVTAED	---	NGMPCFQOL	279
gi 1935019	TVLATSALY	VKETALITPEK	TVVTTED	---	SGMPCFQOL	278
gi 7649151	LMTLVIALTY	VKEKQMSPEI	AKTVAGVDGD	DEDG	AAVLPFFGEI	280
gi 468562	LSLTVLALS	VKEKPMSPDQ	AVDNAEDDTA	SQA	AQPMPPFFGEI	285
CeresGdna:1479767	LITLIALSY	VREKPMSPPEG	SPGEGGDEEE	EEEEI GEAKE	SAPMPFFGEI	250
Consensus	L-LTILAL--	VKEKP---	-----	-----	--A--PFFGEL	300

Figure 13 (continued)

gi|21319 I G A L K D L P K P L L L L V T A L N W I A W F P F L L F D T D W M G K E V Y C G G T  
gi|5823000 F G A L K D L P K P L L L L V T C L N W I A W F P F L L F D T D W M G K E V Y C G G T  
gi|439294 F G A L K E L P R P M W I L L L V T C L N W I A W F P F L L Y D T D W M A K E V F G G Q  
gi|575351 F G A L K D L P R P M W I L L L V T S L N W I A R F P F E L Y D T D W M A K E V Y C G G K  
Lead:CeresClone7894 F G A F K E L K R P M W M L L I V T A L N W I A W F P F L L F D T D W M G R E V Y C G G N S D A T A T  
gi|18091781 F G A F K E L K R P M W M L L I V T A L N W I A W F P F L L F D T D W M G R E V Y C G G N S D A T A S  
gi|415988 F G A L K D L P R P M W V L L L V T A L N W I A W F G F L L F D T D W M G R E V Y C E T Q Q H K  
gi|6120115 F S A L K N L P R P M W F L L L V A C L N W I A W F P F L L F D T D W M G K E V Y C G G T K  
gi|17447420 F G A L K D L P R P M W M L L L V T A L N W I A W F P F L L F D T D W M G K E V Y C G G T K  
gi|68161544 M N A F R Q L R R P M W V L L L V T C L N W I A W F P F L L F D T D W M G K E V Y C G G T  
gi|51863031 M S A F K G L Q R P M R L L L L V T C L N W I A W F P F L L F D T D W M G R E V Y C G G T  
gi|33620334 F G A F R E L K R P M W I L L L V T C L N W I A W F P F L L F D T D W M G K E V Y C G G T  
gi|5230818 S G A F K E L K R P M W I L L L V T C L N W I A W F P F L L F D T D W M G K E V Y E G T  
gi|1935019 S G A F K E L K R P M W I L L L V T C L N W I A W F P F L L F D T D W M G K E V Y C G G T  
gi|7649151 F A A L K N L Q R P M R L L L L V T C L N W I A W F P F L L F D T D W M G K E V Y C G G T  
gi|468562 L G A F K N L K R P M W I L L L V T C L N W I A W F P F L L F D T D W M G R E V Y C G G S S D S S A T  
CeresGdna:1479767 F A A L K N L Q R P M W I L L L V T C L N W A W F P F L L F D T D W M G R E V Y C G G S S D S S A T  
Consensus F G A L K L K R P M W I L L L V T C L N W I A W F P F L L F D T D W M G K E V Y C G G T

gi|21319 V G E G K L Y D Q G V H A G A L G L M I N S V V L G V M S L S J E G L A R M V G G A K R L W G  
gi|5823000 V G E G K A Y D M G V H A G A L G L M I N S V V L G I M S L G I E K L A R L V G G V K R L W G  
gi|439294 V G D A R L Y D L G V R A G A M G L L L Q S V V L G F M S L G V E F L G K K I G G A K R L W G  
gi|575351 V G D G R L Y D L G V H A G A L G L L L N S V V L G F M S L S V E F L G K K I G G V K R L W G  
Lead:CeresClone7894 A A S K K L Y N D G V R A G A L G L M L N A I V L G F M S L G V E M I G R K L G G A K R L W G  
gi|18091781 V A S K K L Y N D C V R A G A L G L M L N A I V L G F M S L G V E M V G R K M G G A K R L W G  
gi|415988 A P E L A V Y N K C V S A G A L G L M L N S I V L G F A S L G V Q Y M A R A L G G V K R L W G  
gi|6120115 V A E G K M Y D R G V R A G A L G L M L N P V V L G F S S L G I Q A I A R G V G G P K R L W G  
gi|17447420 V G E G S L Y D K G V R V G A L G L M L N S V V L G V A S L G V Q V T A R G L G G V K R L W G  
gi|68161544 V G E G P R G R L Y D L G V R A G S L G L M L N S V V L G L M S L G V E F F G R G V G V K R L W G  
gi|51863031 V E R S G G K I V R S G V R A G A L G L M L N S V V M G F T S L G V N V S R G I G G V K R L W G  
gi|33620334 V G E G K A Y D R G V R A G A L G L M L N S V V L G A T S L G V E M L A R G V G V K R L W G  
gi|5230818 V G E G H A Y D M G V R A G A L G L M L N S V V L G A T S L G V D I L A R G V G V K R L W G  
gi|1935019 V G E G H A Y D M C V R E G A L G L M L N S V V L G A T S L G V D I L A R G V G V K R L W G  
gi|7649151 A G E S K L Y D R G V R A G A L G L M L N S V V L G F T S L G V E L M A R K L G V K R L W G  
gi|468562 A E Q L K L Y D R G V R A G A L G L M L N S V V L G F T S L G V E M L A R G V G V K R L W G  
CeresGdna:1479767 A D Q L K M Y D R G V R A G A L G L M L N S V V L G V T S L G V E A L A R G V G V K R L W G  
Consensus V G E G K L Y D G V R A G A L G L M L N S V V L G F T S L G V E L A R G V G V K R L W G

Figure 13 (continued)



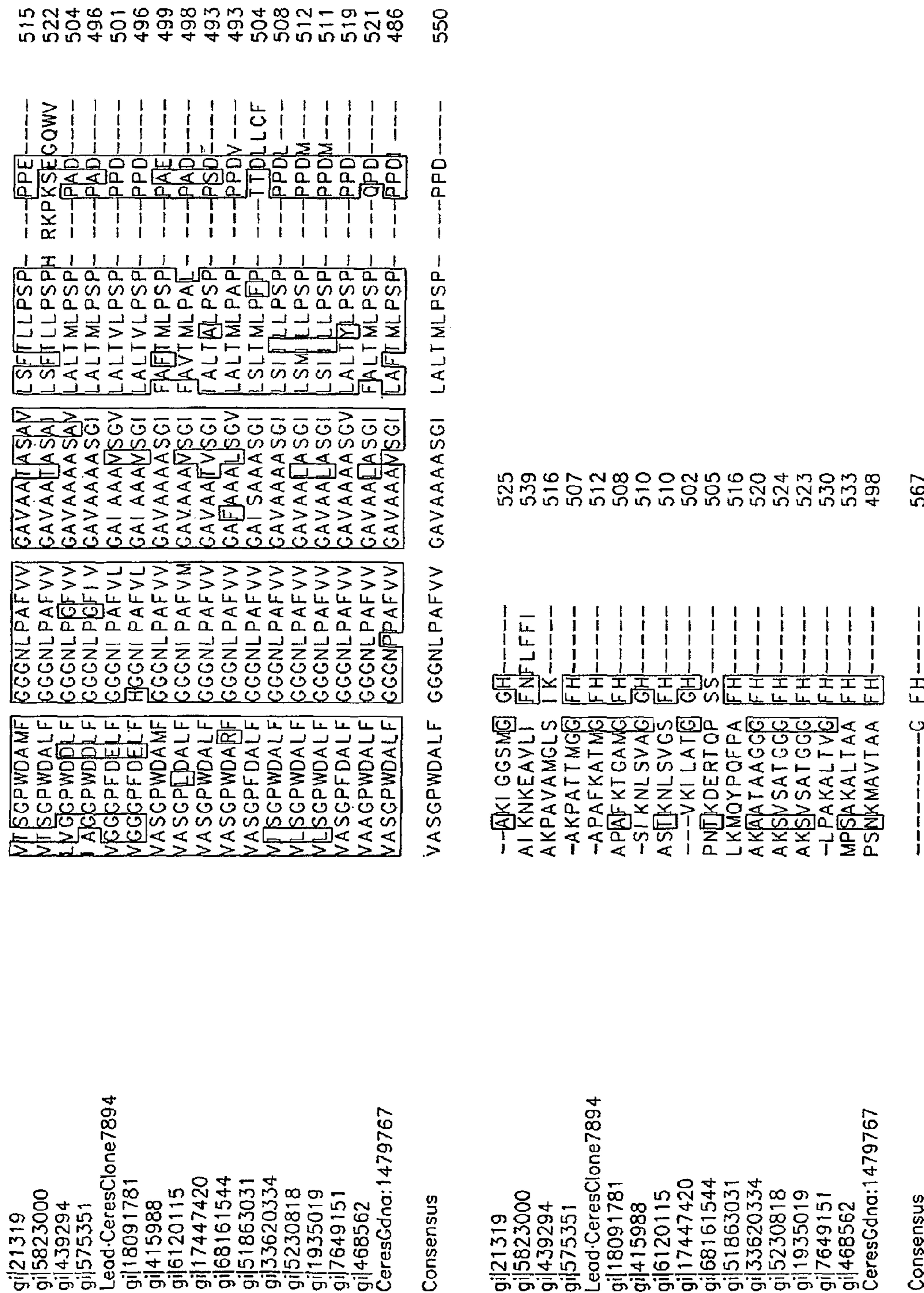


Figure 13 (continued)

**MODULATING PLANT OIL LEVELS****CROSS-REFERENCE TO RELATED APPLICATIONS**

This Application is a National Stage application under 35 U.S.C. §371 of International Application No. PCT/US2007/002214, having an International Filing Date of Jan. 26, 2007, which claims priority under 35 U.S.C. 119(e) to U.S. Provisional Application No. 60/762,422, filed Jan. 26, 2006, and U.S. Provisional Application No. 60/797,077, filed May 1, 2006, all of which are incorporated herein by reference in their entirety.

**BACKGROUND****1. Technical Field**

This document relates to methods and materials involved in modulating (e.g., increasing or decreasing) oil levels in plants. For example, this document provides plants having increased oil levels as well as materials and methods for making plants and plant products having increased oil levels.

**2. Incorporation-by-Reference & Texts**

The material on the accompanying diskette is hereby incorporated by reference into this application. The accompanying compact discs contain one identical file, 11696-204WO1-Sequence.txt, which was created on Jan. 26, 2007. The file named 11696-204WO1-Sequence.txt is 1128 KB. The file can be accessed using Microsoft Word on a computer that uses Windows OS.

**3. Background Information**

Fat, protein, and carbohydrates are nutrients that supply calories to the body. Fat provides nine calories per gram, which is more than twice the number provided by carbohydrates or protein. Dietary fats are composed of fatty acids and glycerol. The glycerol can be converted to glucose by the liver and used as a source of energy. The fatty acids are a good source of energy for many tissues, especially heart and skeletal muscle.

Fatty acids consist of carbon chains of various lengths and a terminal carboxylic acid group. Saturated fatty acids do not contain any double bonds or other functional groups along the chain. A saturated fatty acid has the maximum possible number of hydrogen atoms attached to every carbon atom. Therefore, it is said to be saturated with hydrogen atoms. Eating too much saturated fat is one of the major risk factors for heart disease. Saturated fats are found in animal products such as butter, cheese, whole milk, ice cream, cream, and fatty meats. Saturated fats are also found in some vegetable oils, such as coconut, palm, and palm kernel oils. Most other vegetable oils contain unsaturated fat that helps to lower blood cholesterol if used in place of saturated fat.

Unsaturated fatty acids contain one or more double bonds between carbon atoms and, therefore, two fewer hydrogen atoms per double bond. A fatty acid with a single double bond is called a monounsaturated fatty acid. A fatty acid with two or more double bonds is called a polyunsaturated fatty acid. Polyunsaturated fats are liquid at room temperature, and remain in liquid form even when refrigerated or frozen. Polyunsaturated fats are divided into two families: the omega-3 fats and the omega-6 fats.

The omega-3 family of fatty acids includes alpha-linolenic acid (ALA). ALA is an essential fatty acid that cannot be synthesized in the body and must, therefore, be consumed in the diet. Dietary sources of ALA include canola, flaxseed, flaxseed oil, soybean, and pumpkin seed oil. Omega-3 fatty

acids have been found to reduce the risks of heart problems, lower high blood pressure, and ameliorate autoimmune diseases.

Omega-6 fatty acids are beneficial as well. The omega-6 family of fatty acids includes linoleic acid, which is another essential fatty acid. The body converts linoleic acid to gamma linoleic acid (GLA) and ultimately to prostaglandins, which are hormone-like molecules that help regulate inflammation and blood pressure as well as heart, gastrointestinal, and kidney functions. The main sources of omega-6 fatty acids are vegetable oils such as corn oil and soy oil.

Vegetable oil is fat extracted from plant sources. Vegetable oils are used in cooking, in making margarine and other processed foods, and in producing several non-food items such as soap, cosmetics, medicine, and paint. Since vegetable oils are usually extracted from the seeds of the plant, seed oil yield has a significant impact on the economics of producing many products. Increasing seed oil content may increase the economic return per unit to the seller of the seed in addition to increasing the nutritional value to the consumer of the seed.

**SUMMARY**

This document provides methods and materials related to plants having modulated (e.g., increased or decreased) levels of oil. For example, this document provides transgenic plants and plant cells having increased levels of oil, nucleic acids used to generate transgenic plants and plant cells having increased levels of oil, and methods for making plants and plant cells having increased levels of oil. Such plants and plant cells can be grown to produce, for example, seeds having increased oil content. Increasing the oil content of seeds can increase the nutritional value of the seeds and the yield of oil obtained from the seeds, which may benefit both food consumers and producers.

In one aspect, a method of modulating the level of oil in a plant is provided. The method comprises introducing into a plant cell an isolated nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:81-92, SEQ ID NO:94, SEQ ID NOs:96-107, SEQ ED NO:109, SEQ ID NOs:111-116, SEQ ID NOs:118-134, SEQ ID NOs:136-140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NOs:146-147, SEQ ID NOs:149-150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NOs:156-157, SEQ ID NOs:159-169, SEQ ID NOs:171-172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NOs:178-179, SEQ ID NOs:181-182, SEQ ID NOs:184-191, SEQ ID NOs:193-196, SEQ ID NOs:198-217, SEQ ID NOs:219-220, SEQ ID NOs:222-229, SEQ ID NOs:231-238, SEQ ID NO:240, SEQ ID NOs:242-250, SEQ ID NOs:252-330, SEQ ID NOs:332-333, SEQ ID NO:335, SEQ ID NOs:337-340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346-348, SEQ ID NOs:350-357, SEQ ID NOs:359-366, SEQ ID NO:368, SEQ ID NOs:370-372, SEQ ID NOs:374-376, SEQ ID NOs:378-379, SEQ ID NOs:381-396, SEQ ID NO:398, SEQ ID NOs:502-545, and the consensus sequences set forth in FIGS. 1-13, where a tissue of a plant produced from the plant cell has a difference in the level of oil as compared to the corresponding level in tissue of a control plant that does not comprise the nucleic acid.

In another aspect, a method of modulating the level of oil in a plant is provided. The method comprises introducing into a plant cell an isolated nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:81-82, SEQ ID NOs:

84-85, SEQ ID NO:87, SEQ ID NO:91, SEQ ID NO:94, SEQ ID NOs:96-101, SEQ ID NOs:103-105, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NOs:113-115, SEQ ID NO:118, SEQ ID NOs:120-121, SEQ ID NOs:125-132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NOs:139-140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NOs:146-147, SEQ ID NO:149, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NOs:156-157, SEQ ID NO:159, SEQ ID NOs:162-168, SEQ ID NOs:171-172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:181, SEQ ID NO:184, SEQ ID NOs:193-194, SEQ ID NO:198, SEQ ID NO:207, SEQ ID NOs:215-217, SEQ ID NOs:219-220, SEQ ID NOs:222-223, SEQ ID NOs:231-232, SEQ ID NO:235, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NOs:244-246, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NOs:257-258, SEQ ID NO:261, SEQ ID NOs:265-267, SEQ ID NO:270, SEQ ID NO:279, SEQ ID NOs:287-293, SEQ ID NOs:300-301, SEQ ID NO:304, SEQ ID NOs:307-309, SEQ ID NOs:311-313, SEQ ID NOs:317-318, SEQ ID NO:320, SEQ ID NOs:323-324, SEQ ID NO:327, SEQ ID NO:332, SEQ ID NO:335, SEQ ID NOs:337-340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NOs:346-348, SEQ ID NOs:350-351, SEQ ID NOs:355-357, SEQ ID NOs:359-361, SEQ ID NOs:363-366, SEQ ID NO:368, SEQ ID NOs:370-372, SEQ ID NO:374, SEQ ID NO:378, SEQ ID NO:381, SEQ ID NO:398, SEQ ID NOs:502-505, SEQ ID NOs:507-519, SEQ ID NO:521, SEQ ID NOs:524-527, SEQ ID NOs:529-534, SEQ ID NO:541, SEQ ID NOs:544-545, and the consensus sequences set forth in FIGS. 1-13, where a tissue of a plant produced from the plant cell has a difference in the level of oil as compared to the corresponding level in tissue of a control plant that does not comprise the nucleic acid.

In another aspect, a method of modulating the level of oil in a plant is provided. The method comprises introducing into a plant cell an isolated nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:81-82, SEQ ID NOs:84-85, SEQ ID NO:87, SEQ ID NO:91, SEQ ID NO:94, SEQ ID NOs:96-101, SEQ ID NOs:103-105, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NOs:113-115, SEQ ID NO:118, SEQ ID NOs:120-121, SEQ ID NOs:125-132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NOs:139-140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NOs:146-147, SEQ ID NO:149, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NOs:156-157, SEQ ID NO:159, SEQ ID NOs:162-168, SEQ ID NOs:171-172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:181, SEQ ID NO:184, SEQ ID NOs:193-194, SEQ ID NO:198, SEQ ID NO:207, SEQ ID NOs:215-217, SEQ ID NOs:219-220, SEQ ID NOs:222-223, SEQ ID NOs:231-232, SEQ ID NO:235, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NOs:244-246, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NOs:257-258, SEQ ID NO:261, SEQ ID NOs:265-267, SEQ ID NO:270, SEQ ID NO:279, SEQ ID NOs:287-293, SEQ ID NOs:300-301, SEQ ID NO:304, SEQ ID NOs:307-309, SEQ ID NOs:311-313, SEQ ID NOs:317-318, SEQ ID NO:320, SEQ ID NOs:323-324, SEQ ID NO:327, SEQ ID NO:332, SEQ ID NO:335, SEQ ID NOs:337-340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NOs:346-348, SEQ ID NOs:350-351, SEQ ID NOs:355-357, SEQ ID NOs:359-361, SEQ ID NOs:363-366, SEQ ID NO:368, SEQ ID NOs:370-372, SEQ ID NO:374, SEQ ID NO:378, SEQ ID NO:381, SEQ ID NO:398, SEQ ID NOs:502-505, SEQ ID NOs:507-519, SEQ ID NO:521, SEQ ID NOs:524-527, SEQ ID NOs:529-534, SEQ ID NO:541, and SEQ ID NOs:544-545, where a tissue of a plant produced from the plant cell has a difference in the

level of oil as compared to the corresponding level in tissue of a control plant that does not comprise the nucleic acid.

The sequence identity can be 85 percent or greater, 90 percent or greater, or 95 percent or greater. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:81. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:94. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:111. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:136. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:152. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:159. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:171. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:176. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:178. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:193. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:332. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:342. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:344. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:346. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:359. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:374. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:398. The nucleotide sequence corresponding to a consensus sequence set forth in FIG. 1, FIG. 2, FIG. 3, FIG. 4, FIG. 5, FIG. 6, FIG. 7, FIG. 8, FIG. 9, FIG. 10, FIG. 11, FIG. 12, or FIG. 13. The difference can be an increase in the level of oil. The isolated nucleic acid can be operably linked to a regulatory region. The regulatory region can be a tissue-preferential regulatory region tissue-preferential regulatory region can be a promoter. The regulatory region can be a broadly expressing promoter. The plant can be a dicot. The plant can be a member of the genus *Anacardium*, *Arachis*, *Azadirachta*, *Brassica*, *Cannabis*, *Carthamus*, *Corylus*, *Crambe*, *Cucurbita*, *Glycine*, *Gossypium*, *Helianthus*, *Jatropha*, *Juglans*, *Linum*, *Olea*, *Papaver*, *Persea*, *Prunus*, *Ricinus*, *Sesamum*, *Simmondsia*, or *Vitis*. The plant can be a monocot. The plant can be a member of the genus *Cocos*, *Elaeis*, *Oryza*, or *Zea*. The tissue can be seed tissue.

A method of producing a plant tissue is also provided. The method comprises growing a plant cell comprising an exogenous nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:81-92, SEQ ID NO:94, SEQ ID NOs:96-107, SEQ ID NO:109, SEQ ID NOs:111-116, SEQ ID NOs:118-134, SEQ ID NOs:136-140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NOs:146-147, SEQ ID NOs:149-150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NOs:156-157, SEQ ID NOs:159-169, SEQ ID NOs:171-172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NOs:178-179, SEQ ID



NOs:181-182, SEQ ID NOs:184-191, SEQ ID NOs:193-196, SEQ ID NOs:198-217, SEQ ID NOs:219-220, SEQ ID NOs:222-229, SEQ ID NOs:231-238, SEQ ID NO:240, SEQ ID NOs:242-250, SEQ ID NOs:252-330, SEQ ID NOs:332-333, SEQ ID NO:335, SEQ ID NOs:337-340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346-348, SEQ ID NOs:350-357, SEQ ID NOs:359-366, SEQ ID NO:368, SEQ ID NOs:370-372, SEQ ID NOs:374-376, SEQ ID NOs:378-379, SEQ ID NOs:381-396, SEQ ID NO:398, SEQ ID NOs:502-545, and the consensus sequences set forth in FIGS. 1-13, where the tissue has a difference in the level of oil as compared to the corresponding level in tissue of a control plant that does not comprise the nucleic acid.

In another aspect, a method of producing a plant tissue is provided. The method comprises growing a plant cell comprising an exogenous nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:81-82, SEQ ID NOs:84-85, SEQ ID NO:87, SEQ ID NO:91, SEQ ID NO:94, SEQ ID NOs:96-101, SEQ ID NOs:103-105, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NOs:113-115, SEQ ID NO:118, SEQ ID NOs:120-121, SEQ ID NOs:125-132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NOs:139-140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NOs:146-147, SEQ ID NO:149, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NOs:156-157, SEQ ID NO:159, SEQ ID NOs:162-168, SEQ ID NOs:171-172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:181, SEQ ID NO:184, SEQ ID NOs:193-194, SEQ ID NO:198, SEQ ID NO:207, SEQ ID NOs:215-217, SEQ ID NOs:219-220, SEQ ID NOs:222-223, SEQ ID NOs:231-232, SEQ ID NO:235, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NOs:244-246, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NOs:257-258, SEQ ID NO:261, SEQ ID NOs:265-267, SEQ ID NO:270, SEQ ID NO:279, SEQ ID NOs:287-293, SEQ ID NOs:300-301, SEQ ID NO:304, SEQ ID NOs:307-309, SEQ ID NOs:311-313, SEQ ID NOs:317-318, SEQ ID NO:320, SEQ ID NOs:323-324, SEQ ID NO:327, SEQ ID NO:332, SEQ ID NO:335, SEQ ID NOs:337-340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NOs:346-348, SEQ ID NOs:350-351, SEQ ID NOs:355-357, SEQ ID NOs:359-361, SEQ ID NOs:363-366, SEQ ID NO:368, SEQ ID NOs:370-372, SEQ ID NO:374, SEQ ID NO:378, SEQ ID NO:381, SEQ ID NO:398, SEQ ID NOs:502-505, SEQ ID NOs:507-519, SEQ ID NO:521, SEQ ID NOs:524-527, SEQ ID NOs:529-534, SEQ ID NO:541, SEQ ID NOs:544-545, and the consensus sequences set forth in FIGS. 1-13, where the tissue has a difference in the level of oil as compared to the corresponding level in tissue of a control plant that does not comprise the nucleic acid.

In another aspect, a method of producing a plant tissue is provided. The method comprises growing a plant cell comprising an exogenous nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:81-82, SEQ ID NOs:84-85, SEQ ID NO:87, SEQ ID NO:91, SEQ ID NO:94, SEQ ID NOs:96-101, SEQ ID NOs:103-105, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NOs:113-115, SEQ ID NO:118, SEQ ID NOs:120-121, SEQ ID NOs:125-132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NOs:139-140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NOs:146-147, SEQ ID NO:149, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NOs:156-157, SEQ ID NO:159, SEQ ID NOs:162-168, SEQ ID NOs:171-172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:181, SEQ ID NO:184, SEQ ID NOs:

193-194, SEQ ID NO:198, SEQ ID NO:207, SEQ ID NOs:215-217, SEQ ID NOs:219-220, SEQ ID NOs:222-223, SEQ ID NOs:231-232, SEQ ID NO:235, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NOs:244-246, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NOs:257-258, SEQ ID NO:261, SEQ ID NOs:265-267, SEQ ID NO:270, SEQ ID NO:279, SEQ ID NOs:287-293, SEQ ID NOs:300-301, SEQ ID NO:304, SEQ ID NOs:307-309, SEQ ID NOs:311-313, SEQ ID NOs:317-318, SEQ ID NO:320, SEQ ID NOs:323-324, SEQ ID NO:327, SEQ ID NO:332, SEQ ID NO:335, SEQ ID NOs:337-340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NOs:346-348, SEQ ID NOs:350-351, SEQ ID NOs:355-357, SEQ ID NOs:359-361, SEQ ID NOs:363-366, SEQ ID NO:368, SEQ ID NOs:370-372, SEQ ID NO:374, SEQ ID NO:378, SEQ ID NO:381, SEQ ID NO:398, SEQ ID NOs:502-505, SEQ ID NOs:507-519, SEQ ID NO:521, SEQ ID NOs:524-527, SEQ ID NOs:529-534, SEQ ID NO:541, and SEQ ID NOs:544-545, where the tissue has a difference in the level of oil as compared to the corresponding level in tissue of a control plant that does not comprise the nucleic acid.

The sequence identity can be 85 percent or greater, 90 percent or greater, or 95 percent or greater. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:81. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:94. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:111. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:136. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:152. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:159. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:171. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:176. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:178. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:193. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:332. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:342. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:344. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:346. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:359. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:374. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:398. The difference can be an increase in the level of oil. The exogenous nucleic acid can be operably linked to a regulatory region. The regulatory region can be a tissue-preferential regulatory region. The tissue-preferential regulatory region can be a promoter. The regulatory region can be a broadly expressing promoter. The plant tissue can be dicotyledonous. The plant tissue can be a member of the genus *Anacardium*, *Arachis*, *Azadirachta*, *Brassica*, *Cannabis*, *Carthamus*, *Corylus*, *Crambe*, *Cucurbita*, *Glycine*, *Gossypium*, *Helianthus*, *Jatropha*, *Juglans*, *Linum*, *Olea*, *Papaver*, *Persea*, *Prunus*, *Ricinus*, *Sesamum*,

*Simmondsia*, or *Vitis*. The plant tissue can be monocotyledonous. The plant tissue can be a member of the genus *Cocos*, *Elaeis*, *Oryza*, or *Zea*. The tissue can be seed tissue.

A plant cell is also provided. The plant cell comprises an exogenous nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:81-92, SEQ ID NO:94, SEQ ID NOs:96-107, SEQ ID NO:109, SEQ ID NOs:111-116, SEQ ID NOs:118-134, SEQ ID NOs:136-140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NOs:146-147, SEQ ID NOs:149-150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NOs:156-157, SEQ ID NOs:159-169, SEQ ID NOs:171-172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NOs:178-179, SEQ ID NOs:181-182, SEQ ID NOs:184-191, SEQ ID NOs:193-196, SEQ ID NOs:198-217, SEQ ID NOs:219-220, SEQ ID NOs:222-229, SEQ ID NOs:231-238, SEQ ID NO:240, SEQ ID NOs:242-250, SEQ ID NOs:252-330, SEQ ID NOs:332-333, SEQ ID NO:335, SEQ ID NOs:337-340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346-348, SEQ ID NOs:350-357, SEQ ID NOs:359-366, SEQ ID NO:368, SEQ ID NOs:370-372, SEQ ID NOs:374-376, SEQ ID NOs:378-379, SEQ ID NOs:381-396, SEQ ID NO:398, SEQ ID NOs:502-545, and the consensus sequences set forth in FIGS. 1-13, where a tissue of a plant produced from the plant cell has a difference in the level of oil as compared to the corresponding level in tissue of a control plant that does not comprise the nucleic acid.

In another aspect, a plant cell is provided. The plant cell comprises an exogenous nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:81-82, SEQ ID NOs:84-85, SEQ ID NO:87, SEQ ID NO:91, SEQ ID NO:94, SEQ ID NOs:96-101, SEQ ID NOs:103-105, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NOs:113-115, SEQ ID NO:118, SEQ ID NOs:120-121, SEQ ID NOs:125-132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NOs:139-140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NOs:146-147, SEQ ID NO:149, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NOs:156-157, SEQ ID NO:159, SEQ ID NOs:162-168, SEQ ID NOs:171-172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:181, SEQ ID NO:184, SEQ ID NOs:193-194, SEQ ID NO:198, SEQ ID NO:207, SEQ ID NOs:215-217, SEQ ID NOs:219-220, SEQ ID NOs:222-223, SEQ ID NOs:231-232, SEQ ID NO:235, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NOs:244-246, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NOs:257-258, SEQ ID NO:261, SEQ ID NOs:265-267, SEQ ID NO:270, SEQ ID NO:279, SEQ ID NOs:287-293, SEQ ID NOs:300-301, SEQ ID NO:304, SEQ ID NOs:307-309, SEQ ID NOs:311-313, SEQ ID NOs:317-318, SEQ ID NO:320, SEQ ID NOs:323-324, SEQ ID NO:327, SEQ ID NO:332, SEQ ID NO:335, SEQ ID NOs:337-340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NOs:346-348, SEQ ID NOs:350-351, SEQ ID NOs:355-357, SEQ ID NOs:359-361, SEQ ID NOs:363-366, SEQ ID NO:368, SEQ ID NOs:370-372, SEQ ID NO:374, SEQ ID NO:378, SEQ ID NO:381, SEQ ID NO:398, SEQ ID NOs:502-505, SEQ ID NOs:507-519, SEQ ID NO:521, SEQ ID NOs:524-527, SEQ ID NOs:529-534, SEQ ID NO:541, SEQ ID NOs:544-545, and the consensus sequences set forth in FIGS. 1-13, where a tissue of a plant produced from the plant cell has a difference in the level of oil as compared to the corresponding level in tissue of a control plant that does not comprise the nucleic acid.

In another aspect, a plant cell is provided. The plant cell comprises an exogenous nucleic acid comprising a nucleotide

sequence encoding a polypeptide having 80 percent or greater sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs:81-82, SEQ ID NOs:84-85, SEQ ID NO:87, SEQ ID NO:91, SEQ ID NO:94, SEQ ID NOs:96-101, SEQ ID NOs:103-105, SEQ ID NO:109, SEQ ID NO:111, SEQ ID NOs:113-115, SEQ ID NO:118, SEQ ID NOs:120-121, SEQ ID NOs:125-132, SEQ ID NO:134, SEQ ID NO:136, SEQ ID NOs:139-140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NOs:146-147, SEQ ID NO:149, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NOs:156-157, SEQ ID NO:159, SEQ ID NOs:162-168, SEQ ID NOs:171-172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:181, SEQ ID NO:184, SEQ ID NOs:193-194, SEQ ID NO:198, SEQ ID NO:207, SEQ ID NOs:215-217, SEQ ID NOs:219-220, SEQ ID NOs:222-223, SEQ ID NOs:231-232, SEQ ID NO:235, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NOs:244-246, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:254, SEQ ID NOs:257-258, SEQ ID NO:261, SEQ ID NOs:265-267, SEQ ID NO:270, SEQ ID NO:279, SEQ ID NOs:287-293, SEQ ID NOs:300-301, SEQ ID NO:304, SEQ ID NOs:307-309, SEQ ID NOs:311-313, SEQ ID NOs:317-318, SEQ ID NO:320, SEQ ID NOs:323-324, SEQ ID NO:327, SEQ ID NO:332, SEQ ID NO:335, SEQ ID NOs:337-340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NOs:346-348, SEQ ID NOs:350-351, SEQ ID NOs:355-357, SEQ ID NOs:359-361, SEQ ID NOs:363-366, SEQ ID NO:368, SEQ ID NOs:370-372, SEQ ID NO:374, SEQ ID NO:378, SEQ ID NO:381, SEQ ID NO:398, SEQ ID NOs:502-505, SEQ ID NOs:507-519, SEQ ID NO:521, SEQ ID NOs:524-527, SEQ ID NOs:529-534, SEQ ID NO:541, and SEQ ID NOs:544-545, where a tissue of a plant produced from the plant cell has a difference in the level of oil as compared to the corresponding level in tissue of a control plant that does not comprise the nucleic acid.

The sequence identity can be 85 percent or greater, 90 percent or greater, or 95 percent or greater. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:81. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:94. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:111. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:136. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:152. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:159. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:171. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:176. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:178. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:193. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:332. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:342. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:344. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:346. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:359. The nucleotide sequence can encode a polypeptide comprising an amino acid

sequence corresponding to SEQ ID NO:374. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to SEQ ID NO:398. The nucleotide sequence can encode a polypeptide comprising an amino acid sequence corresponding to a consensus sequence set forth in FIG. 1, FIG. 2, FIG. 3, FIG. 4, FIG. 5, FIG. 6, FIG. 7, FIG. 8, FIG. 9, FIG. 10, FIG. 11, FIG. 12, or FIG. 13. The difference can be an increase in the level of oil. The exogenous nucleic acid can be operably linked to a regulatory region. The regulatory region can be a tissue-preferential regulatory region. The regulatory region can be a promoter. The regulatory region can be a broadly expressing promoter. The plant can be a dicot. The plant can be a member of the genus *Anacardium*, *Arachis*, *Azadirachta*, *Brassica*, *Cannabis*, *Carthamus*, *Corylus*, *Crambe*, *Cucurbita*, *Glycine*, *Gossypium*, *Helianthus*, *Jatropha*, *Juglans*, *Linum*, *Olea*, *Papaver*, *Persea*, *Prunus*, *Ricinus*, *Sesamum*, *Simmondsia*, or *Vitis*. The plant can be a monocot. The plant can be a member of the genus *Cocos*, *Elaeis*, *Oryza*, or *Zea*. The tissue can be seed tissue.

A transgenic plant is also provided. The transgenic plant comprises any of the plant cells described above. Progeny of the transgenic plant are also provided. The progeny have a difference in the level of oil as compared to the level of oil in corresponding control plants that do not comprise the exogenous nucleic acid. Seed and vegetative tissue from the transgenic plant are also provided. In addition, food products and feed products comprising seed and/or vegetative tissue from the transgenic plant are provided. Oil from the seed of the transgenic plant is provided, as is a method of making oil. The method comprises extracting oil from the seed of the transgenic plant.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:95.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:96.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:108.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:109.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:117.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:118.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:141.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:142.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a

nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:143.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:144.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:145.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:146.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:148.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:149.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:153.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:154.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:155.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:156.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:173.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:174.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:180.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:181.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:183.

In another aspect, an isolated nucleic acid is provided. The isolated nucleic acid comprises a nucleotide sequence encoding a polypeptide having 80% or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:184.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:197.







In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:567.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:569.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:570.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:571.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:572.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:573.

In another aspect, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a nucleotide sequence having 95% or greater sequence identity to the nucleotide sequence set forth in SEQ ID NO:575.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Although methods and materials similar or equivalent to those described herein can be used to practice the invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

#### DESCRIPTION OF THE DRAWINGS

FIG. 1 is an alignment of Ceres Clone 41573 (SEQ ID NO:81) with homologous and/or orthologous amino acid sequences Ceres Clone:1560908 (SEQ ID NO:82), gil2739168 (SEQ ID NO:83), Ceres Clone:1314177 (SEQ ID NO:85), Ceres Clone 1371577 (SEQ ID NO:87), and gil50920801 (SEQ ID NO:89). The consensus sequence determined by the alignment is set forth.

FIG. 2 is an alignment of Ceres Clone 25429 (SEQ ID NO:94) with homologous and/or orthologous amino acid sequences Ceres Annot:1488311 (SEQ ID NO:96), Ceres Clone:953928 (SEQ ID NO:97), Ceres Clone:524682 (SEQ ID NO:98), Ceres Clone:1609735 (SEQ ID NO:99), gil42565379 (SEQ ID NO:102), Ceres Clone:426736 (SEQ ID NO:103), and gil24473796 (SEQ ID NO:107). The consensus sequence determined by the alignment is set forth.

FIG. 3 is an alignment of Ceres Clone 5750 (SEQ ID NO:111) with homologous and/or orthologous amino acid sequences gil42362268 (SEQ ID NO:112), gil27435806 (SEQ ID NO:116), gil45935118 (SEQ ID NO:119), Ceres

Clone:1017141 (SEQ ID NO:120), Ceres Clone:1448636 (SEQ ID NO:121), gil50919707 (SEQ ID NO:122), Ceres Clone:947192 (SEQ ID NO:127), and gil55978016 (SEQ ID NO:133). The consensus sequence determined by the alignment is set forth.

FIG. 4 is an alignment of Ceres Clone 218626 (SEQ ID NO:136) with homologous and/or orthologous amino acid sequences gil30409136 (SEQ ID NO:137), Ceres Clone:1571117 (SEQ ID NO:139), and Ceres Annot:1440705 (SEQ ID NO:142). The consensus sequence determined by the alignment is set forth.

FIG. 5 is an alignment of Ceres Clone 121021 (SEQ ID NO:152) with homologous and/or orthologous amino acid sequences Ceres Annot: 1501628 (SEQ ID NO:154) and Ceres Clone:1121512 (SEQ ID NO:157).

FIG. 6 is an alignment of Ceres Clone 158765 (SEQ ID NO:159) with homologous and/or orthologous amino acid sequences gil5669656 (SEQ ID NO:161), Ceres Clone:754061 (SEQ ID NO:162), Ceres Clone:537752 (SEQ ID NO:164), Ceres Clone:282892 (SEQ ID NO:166), and gil50925813 (SEQ ID NO:169).

FIG. 7 is an alignment of Ceres Clone 16403 (SEQ ID NO:171) with homologous and/or orthologous amino acid sequences Ceres Clone:611156 (SEQ ID NO:172) and Ceres Annot:1464944 (SEQ ID NO:174).

FIG. 8 is an alignment of Ceres Clone 28635 (SEQ ID NO:178) with homologous and/or orthologous amino acid sequences gil2463569 (SEQ ID NO:179), Ceres Annot:1514021 (SEQ ID NO:181), gil55710094 (SEQ ID NO:182), gil75859951 (SEQ ID NO:185), gil1449163 (SEQ ID NO:186), gil28208268 (SEQ ID NO:188), gil41224629 (SEQ ID NO:189), gil27475614 (SEQ ID NO:190), and gil38426486 (SEQ ID NO:191).

FIG. 9 is an alignment of Clone 35698 (SEQ ID NO:193) with homologous and/or orthologous amino acid sequences Clone 1380019·T (SEQ ID NO:266), gil441457·T (SEQ ID NO:268), Ceres Annot:1483290·T (SEQ ID NO:270), gil40287554·T (SEQ ID NO:271), gil28569265·T (SEQ ID NO:274), gil22597164·T (SEQ ID NO:278), Clone 617835·T (SEQ ID NO:279), gil5762457·T (SEQ ID NO:281), gil77416935·T (SEQ ID NO:284), gil28569267·T (SEQ ID NO:285), gil50906823·T (SEQ ID NO:295), gil30025160·T (SEQ ID NO:315), gil54402104·T (SEQ ID NO:321), gil52851174·T (SEQ ID NO:326), and gil4100646·T (SEQ ID NO:330).

FIG. 10 is an alignment of Ceres Clone 36412 (SEQ ID NO:332) with homologous and/or orthologous amino acid sequences Ceres Annot:1467033 (SEQ ID NO:335) and Ceres Clone:1641329 (SEQ ID NO:338).

FIG. 11 is an alignment of Ceres Clone 4829 (SEQ ID NO:346) with homologous and/or orthologous amino acid sequences Ceres Annot:1485102 (SEQ ID NO:350), Ceres Clone:1646533 (SEQ ID NO:351), gil29371519 (SEQ ID NO:352), gil45935148 (SEQ ID NO:354), Ceres Clone:359934 (SEQ ID NO:355), and Ceres Clone:839270 (SEQ ID NO:357).

FIG. 12 is an alignment of Ceres Clone 5426 (SEQ ID NO:359) with homologous and/or orthologous amino acid sequences Ceres Clone:1123542 (SEQ ID NO:360), Ceres Annot:1499194 (SEQ ID NO:368), and Ceres Clone:557065 (SEQ ID NO:371).

FIG. 13 is an alignment of Ceres Clone 7894 (SEQ ID NO:374) with homologous and/or orthologous amino acid sequences gil18091781 (SEQ ID NO:375), gil468562 (SEQ ID NO:376), Ceres Annot:1479767 (SEQ ID NO:378), gil7649151 (SEQ ID NO:379), gil33620334 (SEQ ID NO:382), gil439294 (SEQ ID NO:383), gil5230818 (SEQ ID

NO:385), gil17447420 (SEQ ID NO:386), gil1935019 (SEQ ID NO:387), gil51863031 (SEQ ID NO:388), gil575351 (SEQ ID NO:389), gil68161544 (SEQ ID NO:390), gil21319 (SEQ ID NO:392), gil5823000 (SEQ ID NO:393), gil6120115 (SEQ ID NO:395), and gil415988 (SEQ ID NO:396).

#### DETAILED DESCRIPTION

The invention features methods and materials related to modulating (e.g., increasing or decreasing) oil levels in plants. In some embodiments, the plants may also have modulated levels of protein. The methods can include transforming a plant cell with a nucleic acid encoding an oil-modulating polypeptide, wherein expression of the polypeptide results in a modulated level of oil. Plant cells produced using such methods can be grown to produce plants having an increased or decreased oil content. Seeds from such plants may be used to produce, for example, foodstuffs and animal feed having an increased oil content. Producing oil from seeds having an increased oil content can allow manufacturers to increase oil yields.

##### Polypeptides

The term “polypeptide” as used herein refers to a compound of two or more subunit amino acids, amino acid analogs, or other peptidomimetics, regardless of post-translational modification, e.g., phosphorylation or glycosylation. The subunits may be linked by peptide bonds or other bonds such as, for example, ester or ether bonds. The term “amino acid” refers to natural and/or unnatural or synthetic amino acids, including D/L optical isomers. Full-length proteins, analogs, mutants, and fragments thereof are encompassed by this definition.

Polypeptides described herein include oil-modulating polypeptides. Oil-modulating polypeptides can be effective to modulate oil levels when expressed in a plant or plant cell. Modulation of the level of oil can be either an increase or a decrease in the level of oil relative to the corresponding level in a control plant.

An oil-modulating polypeptide can be an enzyme, such as an aldose 1-epimerase, a squalene/phytoene synthase, a ubiquitin-conjugating enzyme, or a cytochrome p450 enzyme. An oil-modulating polypeptide can also be a ribosomal polypeptide, such as a 60S acidic ribosomal protein or a ribosomal protein L28e. An oil-modulating polypeptide can also be a regulatory polypeptide, such as a cyclin-dependent kinase regulatory subunit or a PsbP protein subunit of photosystem II (PSII). An oil-modulating polypeptide can also be a transporter polypeptide, such as a tryptophan/tyrosine permease or a major facilitator superfamily (MFS) transporter. An oil-modulating polypeptide can also be a transcription factor polypeptide, such as an AP2 domain-containing transcription factor polypeptide.

An oil-modulating polypeptide can be an enzyme involved in carbohydrate metabolism, such as an aldose 1-epimerase family polypeptide. Aldose 1-epimerase catalyzes the interconversion of the alpha- and beta-anomers of hexose sugars such as glucose and galactose. SEQ ID NO:81 sets forth the amino acid sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 41573 (SEQ ID NO:80), that is predicted to encode an aldose 1-epimerase family polypeptide. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:81. Alternatively, an oil-modulating polypeptide can be a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:81. For example, an oil-modulating polypeptide can have an amino acid sequence with at least

45% sequence identity, e.g., 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:81.

Amino acid sequences of homologs and/or orthologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:81 are provided in FIG. 1, along with a consensus sequence. A consensus amino acid sequence for such homologs and/or orthologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:81, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 1 provides the amino acid sequences of Ceres Clone 41573 (SEQ ID NO:81), Ceres Clone:1560908 (SEQ ID NO:82), gil2739168 (SEQ ID NO:83), Ceres Clone:1314177 (SEQ ID NO:85), Ceres Clone 1371577 (SEQ ID NO:87), and gil50920801 (SEQ ID NO:89). Other homologs and/or orthologs include Ceres CLONE ID no. 399052 (SEQ ID NO:84), Public GI no. 15824567 (SEQ ID NO:86), Public GI no. 15824565 (SEQ ID NO:88), Public GI no. 50909807 (SEQ ID NO:90), Ceres CLONE ID no. 639223 (SEQ ID NO:91), Public GI no. 37531218 (SEQ ID NO:92), and Ceres Clone:1476735 (SEQ ID NO:502).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:92, SEQ ID NO:502, or the consensus sequence set forth in FIG. 1.

An oil-modulating polypeptide can be a ribosomal polypeptide, such as a 60S acidic ribosomal polypeptide. SEQ ID NO:94 sets forth the amino acid sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 25429 (SEQ ID NO:93), that is predicted to encode a 60S acidic ribosomal polypeptide. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:94. Alternatively, an oil-modulating polypeptide can be a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:94. For example, an oil-modulating polypeptide can have an amino acid sequence with at least 60% sequence identity, e.g., 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:94.

Amino acid sequences of homologs and/or orthologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:94 are provided in FIG. 2, along with a consensus sequence. A consensus amino acid sequence for such homologs and/or orthologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:94, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 2 provides the amino acid sequences of Ceres Clone 25429 (SEQ ID NO:94), Ceres Annot:1488311 (SEQ ID NO:96), Ceres Clone:953928 (SEQ ID NO:97), Ceres Clone:524682 (SEQ ID NO:98), Ceres Clone:1609735 (SEQ ID NO:99), gil42565379 (SEQ ID NO:102), Ceres Clone:426736 (SEQ ID NO:103), and gil24473796 (SEQ ID NO:107). Other homologs and/or orthologs include Ceres CLONE ID no. 949174 (SEQ ID NO:100), Ceres CLONE ID no. 1299820 (SEQ ID NO:101), Ceres CLONE ID no. 1094375 (SEQ ID NO:104), Ceres CLONE ID no. 691062 (SEQ ID NO:105), Public GI no. 47026878 (SEQ ID NO:106), Ceres Annot:



1465437\_PRT (SEQ ID NO:109), Ceres Clone:1798334 (SEQ ID NO:503), Ceres Clone:1886478 (SEQ ID NO:504), Ceres Clone:1727128 (SEQ ID NO:505), and gil730583 (SEQ ID NO:506).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:101, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:503, SEQ ID NO:504, SEQ ID NO:505, SEQ ID NO:506, or the consensus sequence set forth in FIG. 2.

An oil-modulating polypeptide can be a cyclin-dependent kinase regulatory subunit. In eukaryotes, cyclin-dependent protein kinases interact with cyclins to regulate cell cycle progression. The proteins bind to a regulatory subunit, cyclin-dependent kinase regulatory subunit (CKS), which is essential for their function. SEQ ID NO:111 sets forth the amino acid sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 5750 (SEQ ID NO:110), that is predicted to encode a cyclin-dependent kinase regulatory subunit. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:111. Alternatively, an oil-modulating polypeptide can be a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:111. For example, an oil-modulating polypeptide can have an amino acid sequence with at least 75% sequence identity, e.g., 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:111.

Amino acid sequences of homologs and/or orthologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:111 are provided in FIG. 3, along with a consensus sequence. A consensus amino acid sequence for such homologs and/or orthologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:111, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 3 provides the amino acid sequences of Ceres Clone 5750 (SEQ ID NO:111), gil42362268 (SEQ ID NO:112), gil27435806 (SEQ ID NO:116), gil45935118 (SEQ ID NO:119), Ceres Clone:1017141 (SEQ ID NO:120), Ceres Clone: 1448636 (SEQ ID NO:121), gil50919707 (SEQ ID NO:122), Ceres Clone:947192 (SEQ ID NO:127), and gil55978016 (SEQ ID NO:133). Other homologs and/or orthologs include Ceres CLONE ID no. 709027 (SEQ ID NO:113), Ceres CLONE ID no. 853298 (SEQ ID NO:114), Ceres CLONE ID no. 1417425 (SEQ ID NO:115), Ceres Annot:1481954 (SEQ ID NO:118), Public GI no. 40641585 (SEQ ID NO:123), Public GI no. 38566522 (SEQ ID NO:124), Ceres CLONE ID no. 1338131 (SEQ ID NO:125), Ceres CLONE ID no. 300692 (SEQ ID NO:126), Ceres CLONE ID no. 1465004 (SEQ ID NO:128), Ceres CLONE ID no. 1122958 (SEQ ID NO:129), Ceres CLONE ID no. 944737 (SEQ ID NO:130), Ceres CLONE ID no. 217797 (SEQ ID NO:131), Ceres CLONE ID no. 520185 (SEQ ID NO:132), Ceres CLONE ID no. 1436585 (SEQ ID NO:134), Ceres Clone:1777369 (SEQ ID NO:507), Ceres Clone:1744578 (SEQ ID NO:508), Ceres Clone:100008703 (SEQ ID NO:509), and Ceres Clone:1723582 (SEQ ID NO:510).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:112, SEQ

ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:119, SEQ ID NO:120, SEQ ID NO:121, SEQ ID NO:122, SEQ ID NO:123, SEQ ID NO:124, SEQ ID NO:125, SEQ ID NO:126, SEQ ID NO:127, SEQ ID NO:128, SEQ ID NO:129, SEQ ID NO:130, SEQ ID NO:131, SEQ ID NO:132, SEQ ID NO:133, SEQ ID NO:134, SEQ ID NO:507, SEQ ID NO:508, SEQ ID NO:509, SEQ ID NO:510, or the consensus sequence set forth in FIG. 3.

An oil-modulating polypeptide can be a tryptophan/tyrosine permease family polypeptide. Amino acid permeases are integral membrane proteins involved in the transport of amino acids into the cell. SEQ ID NO:136 sets forth the amino acid sequence of a *Zea mays* clone, identified herein as Ceres Clone 218626 (SEQ ID NO:135), that is predicted to encode a tryptophan/tyrosine permease family polypeptide. The tryptophan/tyrosine permease family of proteins is characterized by, inter alia, the presence of several membrane-spanning domains. Members of the tryptophan/tyrosine permease family sometimes have homology to other permeases and transporters. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:136. Alternatively, an oil-modulating polypeptide can be a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:136. For example, an oil-modulating polypeptide can have an amino acid sequence with at least 70% sequence identity, e.g., 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:136.

Amino acid sequences of homologs and/or orthologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:136 are provided in FIG. 4, along with a consensus sequence. A consensus amino acid sequence for such homologs and/or orthologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:136, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 4 provides the amino acid sequences of Ceres Clone 218626 (SEQ ID NO:136), gil30409136 (SEQ ID NO:137), Ceres Clone:1571117 (SEQ ID NO:139), and Ceres Annot:1440705 (SEQ ID NO:142). Other homologs and/or orthologs include Public GI no. 50940751 (SEQ ID NO:138), Ceres CLONE ID no. 424395 (SEQ ID NO:140), Ceres Annot:1493584 (SEQ ID NO:144), Ceres Annot:1463076 (SEQ ID NO:146), Ceres CLONE ID no. 1002421 (SEQ ID NO:147), Ceres Annot:1516369 (SEQ ID NO:149), Public GI no. 30693666 (SEQ ID NO:150), and Ceres Clone:1796001 (SEQ ID NO:511).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:137, SEQ ID NO:138, SEQ ID NO:139, SEQ ID NO:140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NO:146, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:150, SEQ ID NO:511, or the consensus sequence set forth in FIG. 4.

An oil-modulating polypeptide can be a polypeptide that does not have homology to an existing protein family based on Pfam analysis. SEQ ID NO:152 sets forth the amino acid sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 121021 (SEQ ID NO:151), that is predicted to encode a polypeptide that does not have homology to an existing protein family based on Pfam analysis. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:152. Alternatively, an oil-modulating polypeptide can be a homolog having at least 45% sequence

identity, e.g., 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:152.

Amino acid sequences of homologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:152 are provided in FIG. 5, along with a consensus sequence. A consensus amino acid sequence for such homologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:152, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 5 provides the amino acid sequences of Ceres Clone 121021 (SEQ ID NO:152), Ceres Annot: 1501628 (SEQ ID NO:154) and Ceres Clone: 1121512 (SEQ ID NO:157). Other homologs include Ceres Annot:1519046 (SEQ ID NO:156).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:154, SEQ ID NO:156, SEQ ID NO:157, or the consensus sequence set forth in FIG. 5.

SEQ ID NO:159 sets forth the amino acid sequence of another *Arabidopsis* clone, identified herein as Ceres Clone 158765 (SEQ ID NO:158), that is predicted to encode a polypeptide that does not have homology to an existing protein family based on Pfam analysis. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:159. Alternatively, an oil-modulating polypeptide can be a homolog having at least 45% sequence identity, e.g., 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:159.

Amino acid sequences of homologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:159 are provided in FIG. 6, along with a consensus sequence. A consensus amino acid sequence for such homologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:159, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 6 provides the amino acid sequences of Ceres Clone 158765 (SEQ ID NO:159), gil5669656 (SEQ ID NO:161), Ceres Clone:754061 (SEQ ID NO:162), Ceres Clone:537752 (SEQ ID NO:164), Ceres Clone:282892 (SEQ ID NO:166), and gil50925813 (SEQ ID NO:169). Other homologs include Public GI no. 32562996 (SEQ ID NO:160), Ceres CLONE ID no. 1329861 (SEQ ID NO:163), Ceres CLONE ID no. 1322549 (SEQ ID NO:165), Ceres CLONE ID no. 284046 (SEQ ID NO:167), Ceres CLONE ID no. 1388825 (SEQ ID NO:168), and Ceres Clone:1839717 (SEQ ID NO:545).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:160, SEQ ID NO:161, SEQ ID NO:162, SEQ ID NO:163, SEQ ID NO:164, SEQ ID NO:165, SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:545, or the consensus sequence set forth in FIG. 6.

An oil-modulating polypeptide can be a PsbP polypeptide. A PsbP polypeptide is a subunit of photosystem II (PSII) that is reported to be essential for the regulation and stabilization of PSII in higher plants. SEQ ID NO:171 sets forth the amino acid sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 16403 (SEQ ID NO:170), that is predicted to encode a PsbP polypeptide. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:171. Alternatively, an oil-modulating polypeptide can be

a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:171. For example, an oil-modulating polypeptide can have an amino acid sequence with at least 50% sequence identity, e.g., 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:171.

Amino acid sequences of homologs and/or orthologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:171 are provided in FIG. 7, along with a consensus sequence. A consensus amino acid sequence for such homologs and/or orthologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:171, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 7 provides the amino acid sequences of Ceres Clone 16403 (SEQ ID NO:171), Ceres Clone:611156 (SEQ ID NO:172), Ceres Annot:1464944 (SEQ ID NO:174), Ceres Clone:1728680 (SEQ ID NO:530), Ceres Clone:1807796 (SEQ ID NO:531), Ceres Clone:1771837 (SEQ ID NO:532), and Ceres Clone:1773482 (SEQ ID NO:533).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:172, SEQ ID NO:174, SEQ ID NO:530, SEQ ID NO:531, SEQ ID NO:532, SEQ ID NO:533, or the consensus sequence set forth in FIG. 7.

An oil-modulating polypeptide can be a transcription factor that contains an AP2 (APETALA2) DNA-binding domain. AP2 is one of the prototypic members of a family of transcription factors unique to plants, whose distinguishing characteristic is that they contain the so-called AP2 DNA-binding domain. SEQ ID NO:176 sets forth the amino acid sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 19244 (SEQ ID NO:175), that is predicted to encode a transcription factor containing an AP2 DNA-binding domain. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:176. Alternatively, an oil-modulating polypeptide can be a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:176. For example, an oil-modulating polypeptide can have an amino acid sequence with at least 40% sequence identity, e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:176.

An oil-modulating polypeptide can be a squalene/phytoene synthase. Squalene synthase catalyzes the conversion of two molecules of farnesyl diphosphate into squalene, which is the first committed step in the cholesterol biosynthetic pathway. Phytoene synthase catalyzes the conversion of two molecules of geranylgeranyl diphosphate into phytoene, which is the second step in the biosynthesis of carotenoids from isopentenyl diphosphate. SEQ ID NO:178 sets forth the amino acid sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 28635 (SEQ ID NO:177), that is predicted to encode a squalene/phytoene synthase. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:178. Alternatively, an oil-modulating polypeptide can be a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:178. For example, an oil-modulating polypeptide can have an amino acid sequence with at least 70% sequence identity, e.g., 70%,

75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:178.

Amino acid sequences of homologs and/or orthologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:178 are provided in FIG. 8, along with a consensus sequence. A consensus amino acid sequence for such homologs and/or orthologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:178, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 8 provides the amino acid sequences of Ceres Clone 28635 (SEQ ID NO:178), gil2463569 (SEQ ID NO:179), Ceres Annot: 1514021 (SEQ ID NO:181), gil55710094 (SEQ ID NO:182), gil75859951 (SEQ ID NO:185), gil1449163 (SEQ ID NO:186), gil28208268 (SEQ ID NO:188), gil41224629 (SEQ ID NO:189), gil27475614 (SEQ ID NO:190), and gil38426486 (SEQ ID NO:191). Other homologs and/or orthologs include Ceres Annot:1503464 (SEQ ID NO:184) and Public GI no. 1449165 (SEQ ID NO:187), Ceres Clone: 1920025 (SEQ ID NO:534), gil110293133 (SEQ ID NO:535), gil5360655 (SEQ ID NO:536), gil4426953 (SEQ ID NO:537), gil1552717 (SEQ ID NO:538), gil66393825 (SEQ ID NO:539), gil1706774 (SEQ ID NO:540), Ceres Clone:1749989 (SEQ ID NO:541), gil115456049 (SEQ ID NO:542), gil2463567 (SEQ ID NO:543), and Ceres Clone: 706088 (SEQ ID NO:544).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:179, SEQ ID NO:181, SEQ ID NO:182, SEQ ID NO:184, SEQ ID NO:185, SEQ ID NO:186, SEQ ID NO:187, SEQ ID NO:188, SEQ ID NO:189, SEQ ID NO:190, SEQ ID NO:191, SEQ ID NO:534, SEQ ID NO:535, SEQ ID NO:536, SEQ ID NO:537, SEQ ID NO:538, SEQ ID NO:539, SEQ ID NO:540, SEQ ID NO:541, SEQ ID NO:542, SEQ ID NO:543, SEQ ID NO:544, or the consensus sequence set forth in FIG. 8.

An oil-modulating polypeptide can be an ubiquitin-conjugating enzyme. An ubiquitin-conjugating enzyme (E2) is one of at least three enzymes involved in ubiquitination. The E2 enzyme transfers a ubiquitin moiety directly to a substrate, or to a ubiquitin ligase (E3). E2 enzymes are broadly grouped into four classes: class I enzymes contain the catalytic core domain (UBC) having an active site cysteine, class II enzymes possess a UBC and a C-terminal extension, class III enzymes possess a UBC and an N-terminal extension, and class IV enzymes possess a UBC and both N- and C-terminal extensions. These extensions appear to be important for some subfamily function, including E2 localization and protein-protein interactions. In addition, there are proteins with an E2-like fold that are devoid of catalytic activity, but which appear to assist in poly-ubiquitin chain formation. SEQ ID NO:193 sets forth the amino acid sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 35698 (SEQ ID NO:192), that is predicted to encode a ubiquitin-conjugating enzyme. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:193. Alternatively, an oil-modulating polypeptide can be a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:193. For example, an oil-modulating polypeptide can have an amino acid sequence with at least 75% sequence identity, e.g., 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:193.

Amino acid sequences of homologs and/or orthologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:193 are provided in FIG. 9, along with a consensus sequence. A consensus amino acid sequence for such homologs and/or orthologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:193, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 9 provides the amino acid sequences of Ceres Clone 35698 (SEQ ID NO:193), Clone 1380019-T (SEQ ID NO:266), gil441457-T (SEQ ID NO:268), Ceres Annot:1483290-T (SEQ ID NO:270), gil40287554-T (SEQ ID NO:271), gil28569265-T (SEQ ID NO:274), gil22597164-T (SEQ ID NO:278), Clone 617835-T (SEQ ID NO:279), gil5762457-T (SEQ ID NO:281), gil77416935-T (SEQ ID NO:284), gil28569267-T (SEQ ID NO:285), gil50906823-T (SEQ ID NO:295), gil30025160-T (SEQ ID NO:315), gil54402104-T (SEQ ID NO:321), gil52851174-T (SEQ ID NO:326), and gil4100646-T (SEQ ID NO:330). Other homologs and/or orthologs include Ceres CLONE ID no. 1346445 (SEQ ID NO:194), Public GI no. 441457 (SEQ ID NO:195), Public GI no. 19347859 (SEQ ID NO:196), Ceres Annot:1483290 (SEQ ID NO:198), Public GI no. 40287554 (SEQ ID NO:199), Public GI no. 21553796 (SEQ ID NO:200), Public GI no. 28569271 (SEQ ID NO:201), Public GI no. 28569265 (SEQ ID NO:202), Public GI no. 66354420 (SEQ ID NO:203), Public GI no. 21280893 (SEQ ID NO:204), Public GI no. 21554343 (SEQ ID NO:205), Public GI no. 22597164 (SEQ ID NO:206), Ceres CLONE ID no. 617835 (SEQ ID NO:207), Public GI no. 30693871 (SEQ ID NO:208), Public GI no. 5762457 (SEQ ID NO:209), Public GI no. 464981 (SEQ ID NO:210), Public GI no. 456568 (SEQ ID NO:211), Public GI no. 77416935 (SEQ ID NO:212), Public GI no. 28569267 (SEQ ID NO:213), Public GI no. 28569261 (SEQ ID NO:214), Ceres CLONE ID no. 39130 (SEQ ID NO:215), Ceres CLONE ID no. 16865 (SEQ ID NO:216), Ceres CLONE ID no. 575067 (SEQ ID NO:217), Ceres Annot: 1467392 (SEQ ID NO:219), Ceres CLONE ID no. 25162 (SEQ ID NO:220), Ceres Annot:1529647 (SEQ ID NO:222), Ceres CLONE ID no. 1405728 (SEQ ID NO:223), Public GI no. 54288726 (SEQ ID NO:224), Public GI no. 50906823 (SEQ ID NO:225), Public GI no. 83306206 (SEQ ID NO:226), Public GI no. 20152203 (SEQ ID NO:227), Public GI no. 40287568 (SEQ ID NO:228), Public GI no. 50929483 (SEQ ID NO:229), Ceres Annot: 1450556 (SEQ ID NO:231), Ceres CLONE ID no. 1031152 (SEQ ID NO:232), Public GI no. 52548244 (SEQ ID NO:233), Public GI no. 82621144 (SEQ ID NO:234), Ceres CLONE ID no. 878043 (SEQ ID NO:235), Public GI no. 34909292 (SEQ ID NO:236), Public GI no. 2668744 (SEQ ID NO:237), Ceres CLONE ID no. 511132 (SEQ ID NO:238), Ceres Annot:1533930 (SEQ ID NO:240), Ceres Annot:1495171 (SEQ ID NO:242), Public GI no. 20086317 (SEQ ID NO:243), Ceres CLONE ID no. 10022 (SEQ ID NO:244), Ceres CLONE ID no. 12547 (SEQ ID NO:245), Ceres CLONE ID no. 27679 (SEQ ID NO:246), Public GI no. 20259611 (SEQ ID NO:247), Public GI no. 30025160 (SEQ ID NO:248), Public GI no. 50904839 (SEQ ID NO:249), Ceres CLONE ID no. 1063753 (SEQ ID NO:250), Ceres Annot: 1533218 (SEQ ID NO:252), Public GI no. 297880 (SEQ ID NO:253), Ceres CLONE ID no. 1357060 (SEQ ID NO:254), Public GI no. 54402104 (SEQ ID NO:255), Public GI no. 50725323 (SEQ ID NO:256), Ceres CLONE ID no. 376667 (SEQ ID NO:257), Ceres CLONE ID no. 256705 (SEQ ID NO:258), Public GI no. 20259629 (SEQ ID NO:259), Public GI no. 52851174 (SEQ ID NO:260), Ceres CLONE ID no. 1061097 (SEQ ID NO:261), Public GI

no. 1373001 (SEQ ID NO:262), Public GI no. 66354468 (SEQ ID NO:263), Public GI no. 4100646 (SEQ ID NO:264), Ceres CLONE ID no. 1380019 (SEQ ID NO:265), Ceres CLONE ID no. 1346445\_T (SEQ ID NO:267), Public GI no. 19347859\_T (SEQ ID NO:269), Public GI no. 21553796\_T (SEQ ID NO:272), Public GI no. 28569271\_T (SEQ ID NO:273), Public GI no. 66354420\_T (SEQ ID NO:275), Public GI no. 21280893\_T (SEQ ID NO:276), Public GI no. 21554343\_T (SEQ ID NO:277), Public GI no. 30693871\_T (SEQ ID NO:280), Public GI no. 464981\_T (SEQ ID NO:282), Public GI no. 456568\_T (SEQ ID NO:283), Public GI no. 28569261\_T (SEQ ID NO:286), Ceres CLONE ID no. 39130\_T (SEQ ID NO:287), Ceres CLONE ID no. 16865\_T (SEQ ID NO:288), Ceres CLONE ID no. 575067\_T (SEQ ID NO:289), Ceres Annot:1467392\_T (SEQ ID NO:290), Ceres CLONE ID no. 25162\_T (SEQ ID NO:291), Ceres Annot:1529647\_T (SEQ ID NO:292), Ceres CLONE ID no. 1405728\_T (SEQ ID NO:293), Public GI no. 54288726\_T (SEQ ID NO:294), Public GI no. 83306206\_T (SEQ ID NO:296), Public GI no. 20152203\_T (SEQ ID NO:297), Public GI no. 40287568\_T (SEQ ID NO:298), Public GI no. 50929483\_T (SEQ ID NO:299), Ceres Annot:1450556\_T (SEQ ID NO:300), Ceres CLONE ID no. 1031152\_T (SEQ ID NO:301), Public GI no. 52548244\_T (SEQ ID NO:302), Public GI no. 82621144\_T (SEQ ID NO:303), Ceres CLONE ID no. 878043\_T (SEQ ID NO:304), Public GI no. 34909292\_T (SEQ ID NO:305), Public GI no. 2668744\_T (SEQ ID NO:306), Ceres CLONE ID no. 511132\_T (SEQ ID NO:307), Ceres Annot:1533930\_T (SEQ ID NO:308), Ceres Annot:1495171\_T (SEQ ID NO:309), Public GI no. 20086317\_T (SEQ ID NO:310), Ceres CLONE ID no. 10022\_T (SEQ ID NO:311), Ceres CLONE ID no. 12547\_T (SEQ ID NO:312), Ceres CLONE ID no. 27679\_T (SEQ ID NO:313), Public GI no. 20259611\_T (SEQ ID NO:314), Public GI no. 50904839\_T (SEQ ID NO:316), Ceres CLONE ID no. 1063753\_T (SEQ ID NO:317), Ceres Annot:1533218\_T (SEQ ID NO:318), Public GI no. 297880\_T (SEQ ID NO:319), Ceres CLONE ID no. 1357060\_T (SEQ ID NO:320), Public GI no. 50725323\_T (SEQ ID NO:322), Ceres CLONE ID no. 376667\_T (SEQ ID NO:323), Ceres CLONE ID no. 256705\_T (SEQ ID NO:324), Public GI no. 20259629\_T (SEQ ID NO:325), Ceres CLONE ID no. 1061097\_T (SEQ ID NO:327), Public GI no. 1373001\_T (SEQ ID NO:328), and Public GI no. 66354468\_T (SEQ ID NO:329).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:194, SEQ ID NO:195, SEQ ID NO:196, SEQ ID NO:198, SEQ ID NO:199, SEQ ID NO:200, SEQ ID NO:201, SEQ ID NO:202, SEQ ID NO:203, SEQ ID NO:204, SEQ ID NO:205, SEQ ID NO:206, SEQ ID NO:207, SEQ ID NO:208, SEQ ID NO:209, SEQ ID NO:210, SEQ ID NO:211, SEQ ID NO:212, SEQ ID NO:213, SEQ ID NO:214, SEQ ID NO:215, SEQ ID NO:216, SEQ ID NO:217, SEQ ID NO:219, SEQ ID NO:220, SEQ ID NO:222, SEQ ID NO:223, SEQ ID NO:224, SEQ ID NO:225, SEQ ID NO:226, SEQ ID NO:227, SEQ ID NO:228, SEQ ID NO:229, SEQ ID NO:231, SEQ ID NO:232, SEQ ID NO:233, SEQ ID NO:234, SEQ ID NO:235, SEQ ID NO:236, SEQ ID NO:237, SEQ ID NO:238, SEQ ID NO:240, SEQ ID NO:242, SEQ ID NO:243, SEQ ID NO:244, SEQ ID NO:245, SEQ ID NO:246, SEQ ID NO:247, SEQ ID NO:248, SEQ ID NO:249, SEQ ID NO:250, SEQ ID NO:252, SEQ ID NO:253, SEQ ID NO:254, SEQ ID NO:255, SEQ ID

NO:256, SEQ ID NO:257, SEQ ID NO:258, SEQ ID NO:259, SEQ ID NO:260, SEQ ID NO:261, SEQ ID NO:262, SEQ ID NO:263, SEQ ID NO:264, SEQ ID NO:265, SEQ ID NO:266, SEQ ID NO:267, SEQ ID NO:268, SEQ ID NO:269, SEQ ID NO:270, SEQ ID NO:271, SEQ ID NO:272, SEQ ID NO:273, SEQ ID NO:274, SEQ ID NO:275, SEQ ID NO:276, SEQ ID NO:277, SEQ ID NO:278, SEQ ID NO:279, SEQ ID NO:280, SEQ ID NO:281, SEQ ID NO:282, SEQ ID NO:283, SEQ ID NO:284, SEQ ID NO:285, SEQ ID NO:286, SEQ ID NO:287, SEQ ID NO:288, SEQ ID NO:289, SEQ ID NO:290, SEQ ID NO:291, SEQ ID NO:292, SEQ ID NO:293, SEQ ID NO:294, SEQ ID NO:295, SEQ ID NO:296, SEQ ID NO:297, SEQ ID NO:298, SEQ ID NO:299, SEQ ID NO:300, SEQ ID NO:301, SEQ ID NO:302, SEQ ID NO:303, SEQ ID NO:304, SEQ ID NO:305, SEQ ID NO:306, SEQ ID NO:307, SEQ ID NO:308, SEQ ID NO:309, SEQ ID NO:310, SEQ ID NO:311, SEQ ID NO:312, SEQ ID NO:313, SEQ ID NO:314, SEQ ID NO:315, SEQ ID NO:316, SEQ ID NO:317, SEQ ID NO:318, SEQ ID NO:319, SEQ ID NO:320, SEQ ID NO:321, SEQ ID NO:322, SEQ ID NO:323, SEQ ID NO:324, SEQ ID NO:325, SEQ ID NO:326, SEQ ID NO:327, SEQ ID NO:328, SEQ ID NO:329, SEQ ID NO:330, or the consensus sequence set forth in FIG. 9.

SEQ ID NO:332 sets forth the amino acid sequence of another *Arabidopsis* clone, identified herein as Ceres Clone 36412 (SEQ ID NO:331), that is predicted to encode a polypeptide that does not have homology to an existing protein family based on Pfam analysis. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:332. Alternatively, an oil-modulating polypeptide can be a variant having at least 45% sequence identity, e.g., 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:332.

Amino acid sequences of homologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:332 are provided in FIG. 10, along with a consensus sequence. A consensus amino acid sequence for such homologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:332, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 10 provides the amino acid sequences of Ceres Clone 36412 (SEQ ID NO:332), Ceres Annot:1467033 (SEQ ID NO:335) and Ceres Clone:1641329 (SEQ ID NO:338). Other homologs include Public GI no. 3152583 (SEQ ID NO:333), Ceres Annot:1536919 (SEQ ID NO:337), Ceres CLONE ID no. 1650419 (SEQ ID NO:339), and Ceres CLONE ID no. 597699 (SEQ ID NO:340).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:333, SEQ ID NO:335, SEQ ID NO:337, SEQ ID NO:338, SEQ ID NO:339, SEQ ID NO:340, or the consensus sequence set forth in FIG. 10.

SEQ ID NO:342 sets forth the amino acid sequence of another *Arabidopsis* clone, identified herein as Ceres Clone 368 (SEQ ID NO:341), that is predicted to encode a polypeptide that does not have homology to an existing protein family based on Pfam analysis. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:342. Alternatively, an oil-modulating polypeptide can be a homolog having at least 40% sequence identity, e.g., 40%,

45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:342

An oil-modulating polypeptide can be a cytochrome p450 polypeptide. Cytochrome p450 enzymes are haem-thiolate polypeptides involved in the oxidative modification of various compounds. SEQ ID NO:344 sets forth the amino acid sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 41046 (SEQ ID NO:343), that is predicted to encode a cytochrome p450 enzyme. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:344. Alternatively, an oil-modulating polypeptide can be a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:344. For example, an oil-modulating polypeptide can have an amino acid sequence with at least 40% sequence identity, e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:344.

An oil-modulating polypeptide can have a DUF662 domain characteristic of a family of hypothetical eukaryotic proteins. SEQ ID NO:346 sets forth the amino acid sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 4829 (SEQ ID NO:345), that is predicted to encode a eukaryotic polypeptide. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:346. Alternatively, an oil-modulating polypeptide can be a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:346. For example, an oil-modulating polypeptide can have an amino acid sequence with at least 50% sequence identity, e.g., 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:346.

Amino acid sequences of homologs and/or orthologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:346 are provided in FIG. 11, along with a consensus sequence. A consensus amino acid sequence for such homologs and/or orthologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:346, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 11 provides the amino acid sequences of Ceres Clone 4829 (SEQ ID NO:346), Ceres Annot:1485102 (SEQ ID NO:350), Ceres Clone:1646533 (SEQ ID NO:351), gil29371519 (SEQ ID NO:352), gil45935148 (SEQ ID NO:354), Ceres Clone:359934 (SEQ ID NO:355), and Ceres Clone:839270 (SEQ ID NO:357). Other homologs and/or orthologs include Ceres CLONE ID no. 24885 (SEQ ID NO:347), Ceres CLONE ID no. 27878 (SEQ ID NO:348), Public GI no. 38347602 (SEQ ID NO:353), Ceres CLONE ID no. 294598 (SEQ ID NO:356), Ceres Clone:1836904 (SEQ ID NO:525), Ceres Clone:1932013 (SEQ ID NO:526), and Ceres Clone:1768109 (SEQ ID NO:527).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:347, SEQ ID NO:348, SEQ ID NO:350, SEQ ID NO:351, SEQ ID NO:352, SEQ ID NO:353, SEQ ID NO:354, SEQ ID NO:355, SEQ ID NO:356, SEQ ID NO:357, SEQ ID NO:525, SEQ ID NO:526, SEQ ID NO:527, or the consensus sequence set forth in FIG. 11.

An oil-modulating polypeptide can be a ribosomal protein L28e. Ribosomal protein L28e forms part of the 60S ribosomal subunit. SEQ ID NO:359 sets forth the amino acid

sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 5426 (SEQ ID NO:358), that is predicted to encode a ribosomal protein L28e. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:359. Alternatively, an oil-modulating polypeptide can be a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:359. For example, an oil-modulating polypeptide can have an amino acid sequence with at least 65% sequence identity, e.g., 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:359.

Amino acid sequences of homologs and/or orthologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:359 are provided in FIG. 12, along with a consensus sequence. A consensus amino acid sequence for such homologs and/or orthologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:359, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 12 provides the amino acid sequences of Ceres Clone 5426 (SEQ ID NO:359), Ceres Clone:1123542 (SEQ ID NO:360), Ceres Annot:1499194 (SEQ ID NO:368), and Ceres Clone:557065 (SEQ ID NO:371). Other homologs and/or orthologs include Ceres CLONE ID no. 9083 (SEQ ID NO:361), Public GI no. 79322493 (SEQ ID NO:362), Ceres CLONE ID no. 265408 (SEQ ID NO:363), Ceres CLONE ID no. 32164 (SEQ ID NO:364), Ceres CLONE ID no. 1068047 (SEQ ID NO:365), Ceres CLONE ID no. 965035 (SEQ ID NO:366), Ceres Annot:1439584 (SEQ ID NO:370), Ceres CLONE ID no. 465060 (SEQ ID NO:372), Ceres Clone:1458107 (SEQ ID NO:512), Ceres Clone:1932511 (SEQ ID NO:513), Ceres Clone:1850967 (SEQ ID NO:514), Ceres Clone:1835707 (SEQ ID NO:515), Ceres Clone:1727213 (SEQ ID NO:516), Ceres Clone:1767577 (SEQ ID NO:517), Ceres Clone:1712104 (SEQ ID NO:518), Ceres Clone:1778377 (SEQ ID NO:519), gil76573317 (SEQ ID NO:520), Ceres Clone:1787980 (SEQ ID NO:521), gil115465181 (SEQ ID NO:522), gil48716267 (SEQ ID NO:523), and Ceres Clone:575833 (SEQ ID NO:524).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:360, SEQ ID NO:361, SEQ ID NO:362, SEQ ID NO:363, SEQ ID NO:364, SEQ ID NO:365, SEQ ID NO:366, SEQ ID NO:368, SEQ ID NO:370, SEQ ID NO:371, SEQ ID NO:372, SEQ ID NO:512, SEQ ID NO:513, SEQ ID NO:514, SEQ ID NO:515, SEQ ID NO:516, SEQ ID NO:517, SEQ ID NO:518, SEQ ID NO:519, SEQ ID NO:520, SEQ ID NO:521, SEQ ID NO:522, SEQ ID NO:523, SEQ ID NO:524, or the consensus sequence set forth in FIG. 12.

An oil-modulating polypeptide can be a Major Facilitator Superfamily (MFS) transporter. MFS transporters are single-polypeptide secondary carriers capable of transporting small solutes in response to chemiosmotic ion gradients. SEQ ID NO:374 sets forth the amino acid sequence of an *Arabidopsis* clone, identified herein as Ceres Clone 7894 (SEQ ID NO:373), that is predicted to encode an MFS transporter polypeptide. The MFS family of proteins is characterized by the presence of 12 membrane-spanning domains and a conserved MFS-specific motif between membrane-spanning domains 2 and 3. Members of the MFS protein family have been further classified into a number of sub-families. An oil-modulating polypeptide can comprise the amino acid

sequence set forth in SEQ ID NO:374. Alternatively, an oil-modulating polypeptide can be a homolog, ortholog, or variant of the polypeptide having the amino acid sequence set forth in SEQ ID NO:374. For example, an oil-modulating polypeptide can have an amino acid sequence with at least 60% sequence identity, e.g., 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:374. Such oil-modulating polypeptides may have a domain that is a sub-family of MFS transporter polypeptides.

Amino acid sequences of homologs and/or orthologs of the polypeptide having the amino acid sequence set forth in SEQ ID NO:374 are provided in FIG. 13, along with a consensus sequence. A consensus amino acid sequence for such homologs and/or orthologs was determined by aligning amino acid sequences, e.g., amino acid sequences related to SEQ ID NO:374, from a variety of species and determining the most common amino acid or type of amino acid at each position. For example, the alignment in FIG. 13 provides the amino acid sequences of Ceres Clone 7894 (SEQ ID NO:374), gi118091781 (SEQ ID NO:375), gi1468562 (SEQ ID NO:376), Ceres Annot:1479767 (SEQ ID NO:378), gi17649151 (SEQ ID NO:379), gi133620334 (SEQ ID NO:382), gi1439294 (SEQ ID NO:383), gi15230818 (SEQ ID NO:385), gi117447420 (SEQ ID NO:386), gi11935019 (SEQ ID NO:387), gi151863031 (SEQ ID NO:388), gi1575351 (SEQ ID NO:389), gi168161544 (SEQ ID NO:390), gi121319 (SEQ ID NO:392), gi15823000 (SEQ ID NO:393), gi16120115 (SEQ ID NO:395), and gi1415988 (SEQ ID NO:396). Other homologs and/or orthologs include Ceres Annot:1486712 (SEQ ID NO:381), Public GI no. 77153413 (SEQ ID NO:384), Public GI no. 6434833 (SEQ ID NO:391), Public GI no. 633172 (SEQ ID NO:394), gi116008246 (SEQ ID NO:528), and Ceres Clone:1925996 (SEQ ID NO:529).

In some cases, an oil-modulating polypeptide includes a polypeptide having at least 80% sequence identity, e.g., 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to an amino acid sequence corresponding to SEQ ID NO:375, SEQ ID NO:376, SEQ ID NO:378, SEQ ID NO:379, SEQ ID NO:381, SEQ ID NO:382, SEQ ID NO:383, SEQ ID NO:384, SEQ ID NO:385, SEQ ID NO:386, SEQ ID NO:387, SEQ ID NO:388, SEQ ID NO:389, SEQ ID NO:390, SEQ ID NO:391, SEQ ID NO:392, SEQ ID NO:393, SEQ ID NO:394, SEQ ID NO:395, SEQ ID NO:396, SEQ ID NO:528, SEQ ID NO:529, or the consensus sequence set forth in FIG. 13.

SEQ ID NO:398 sets forth the amino acid sequence of another *Arabidopsis* clone, identified herein as Ceres Clone 8161 (SEQ ID NO:397), that is predicted to encode a polypeptide that does not have homology to an existing protein family based on Pfam analysis. An oil-modulating polypeptide can comprise the amino acid sequence set forth in SEQ ID NO:398. Alternatively, an oil-modulating polypeptide can be a homolog having at least 40% sequence identity, e.g., 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the amino acid sequence set forth in SEQ ID NO:398.

An oil-modulating polypeptide encoded by a recombinant nucleic acid can be a native oil-modulating polypeptide, i.e., one or more additional copies of the coding sequence for an oil-modulating polypeptide that is naturally present in the cell. Alternatively, an oil-modulating polypeptide can be heterologous to the cell, e.g., a transgenic *Lycopersicon* plant can contain the coding sequence for a transporter polypeptide from a *Glycine* plant.

An oil-modulating polypeptide can include additional amino acids that are not involved in oil modulation, and thus

can be longer than would otherwise be the case. For example, an oil-modulating polypeptide can include an amino acid sequence that functions as a reporter. Such an oil-modulating polypeptide can be a fusion protein in which a green fluorescent protein (GFP) polypeptide is fused to, e.g., SEQ ID NO:136, or in which a yellow fluorescent protein (YFP) polypeptide is fused to, e.g., SEQ ID NO:81. In some embodiments, an oil-modulating polypeptide includes a purification tag, a chloroplast transit peptide, a mitochondrial transit peptide, or a leader sequence added to the amino or carboxy terminus.

Oil-modulating polypeptide candidates suitable for use in the invention can be identified by analysis of nucleotide and polypeptide sequence alignments. For example, performing a query on a database of nucleotide or polypeptide sequences can identify homologs and/or orthologs of oil-modulating polypeptides. Sequence analysis can involve BLAST, Reciprocal BLAST, or PSI-BLAST analysis of nonredundant databases using known oil-modulating polypeptide amino acid sequences. Those polypeptides in the database that have greater than 40% sequence identity can be identified as candidates for further evaluation for suitability as an oil-modulating polypeptide. Amino acid sequence similarity allows for conservative amino acid substitutions, such as substitution of one hydrophobic residue for another or substitution of one polar residue for another. If desired, manual inspection of such candidates can be carried out in order to narrow the number of candidates to be further evaluated. Manual inspection can be performed by selecting those candidates that appear to have domains suspected of being present in oil-modulating polypeptides, e.g., conserved functional domains.

The identification of conserved regions in a template or subject polypeptide can facilitate production of variants of wild type oil-modulating polypeptides. Conserved regions can be identified by locating a region within the primary amino acid sequence of a template polypeptide that is a repeated sequence, forms some secondary structure (e.g., helices and beta sheets), establishes positively or negatively charged domains, or represents a protein motif or domain. See, e.g., the Pfam web site describing consensus sequences for a variety of protein motifs and domains on the World Wide Web at [sanger.ac.uk/Software/Pfam/](http://sanger.ac.uk/Software/Pfam/) and [pfam.janelia.org/](http://pfam.janelia.org/). A description of the information included at the Pfam database is described in Sonnhammer et al., *Nucl. Acids Res.*, 26:320-322 (1998); Sonnhammer et al., *Proteins*, 28:405-420 (1997); and Bateman et al., *Nucl. Acids Res.*, 27:260-262 (1999).

Conserved regions also can be determined by aligning sequences of the same or related polypeptides from closely related species. Closely related species preferably are from the same family. In some embodiments, alignment of sequences from two different species is adequate. For example, sequences from *Arabidopsis* and *Zea mays* can be used to identify one or more conserved regions.

Typically, polypeptides that exhibit at least about 40% amino acid sequence identity are useful to identify conserved regions. Conserved regions of related polypeptides can exhibit at least 45% amino acid sequence identity (e.g., at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% amino acid sequence identity). In some embodiments, a conserved region of target and template polypeptides exhibit at least 92%, 94%, 96%, 98%, or 99% amino acid sequence identity. Amino acid sequence identity can be deduced from amino acid or nucleotide sequences. In certain cases, highly conserved domains have been identified within oil-modulat-

ing polypeptides. These conserved regions can be useful in identifying functionally similar (orthologous) oil-modulating polypeptides.

In some instances, suitable oil-modulating polypeptides can be synthesized on the basis of consensus functional domains and/or conserved regions in polypeptides that are homologous oil-modulating polypeptides. Domains are groups of substantially contiguous amino acids in a polypeptide that can be used to characterize protein families and/or parts of proteins. Such domains have a “fingerprint” or “signature” that can comprise conserved (1) primary sequence, (2) secondary structure, and/or (3) three-dimensional conformation. Generally, domains are correlated with specific in vitro and/or in vivo activities. A domain can have a length of from 10 amino acids to 400 amino acids, e.g., 10 to 50 amino acids, or 25 to 100 amino acids, or 35 to 65 amino acids, or 35 to 55 amino acids, or 45 to 60 amino acids, or 200 to 300 amino acids, or 300 to 400 amino acids.

Representative homologs and/or orthologs of oil-modulating polypeptides are shown in FIGS. 1-13. Each Figure represents an alignment of the amino acid sequence of an oil-modulating polypeptide with the amino acid sequences of corresponding homologs and/or orthologs. Amino acid sequences of oil-modulating polypeptides and their corresponding homologs and/or orthologs have been aligned to identify conserved amino acids and to determine consensus sequences that contain frequently occurring amino acid residues at particular positions in the aligned sequences, as shown in FIGS. 1-13. A dash in an aligned sequence represents a gap, i.e., a lack of an amino acid at that position. Identical amino acids or conserved amino acid substitutions among aligned sequences are identified by boxes.

Each consensus sequence is comprised of conserved regions. Each conserved region contains a sequence of contiguous amino acid residues. A dash in a consensus sequence indicates that the consensus sequence either lacks an amino acid at that position or includes an amino acid at that position. If an amino acid is present, the residue at that position corresponds to one found in any aligned sequence at that position.

Useful polypeptides can be constructed based on the consensus sequence in FIG. 1, FIG. 2, FIG. 3, FIG. 4, FIG. 5, FIG. 6, FIG. 7, FIG. 8, FIG. 9, FIG. 10, FIG. 11, FIG. 12, or FIG. 13. Such a polypeptide includes the conserved regions in the selected consensus sequence, arranged in the order depicted in the Figure from amino-terminal end to carboxy-terminal end. Such a polypeptide may also include zero, one, or more than one amino acid in positions marked by dashes. When no amino acids are present at positions marked by dashes, the length of such a polypeptide is the sum of the amino acid residues in all conserved regions. When amino acids are present at all positions marked by dashes, such a polypeptide has a length that is the sum of the amino acid residues in all conserved regions and all dashes.

Consensus domains and conserved regions can be identified by homologous polypeptide sequence analysis as described above. The suitability of polypeptides for use as oil-modulating polypeptides can be evaluated by functional complementation studies.

#### Nucleic Acids

Isolated nucleic acids are provided herein. The terms “nucleic acid” and “polynucleotide” are used interchangeably herein, and refer to both RNA and DNA, including cDNA, genomic DNA, synthetic DNA, and DNA (or RNA) containing nucleic acid analogs. Polynucleotides can have any three-dimensional structure. A nucleic acid can be double-stranded or single-stranded (i.e., a sense strand or an antisense strand). Non-limiting examples of polynucleotides

include genes, gene fragments, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, siRNA, micro-RNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers, as well as nucleic acid analogs.

Nucleic acids described herein include oil-modulating nucleic acids. Oil-modulating nucleic acids can be effective to modulate protein levels when transcribed in a plant or plant cell. A oil-modulating nucleic acid can comprise the nucleotide sequence set forth in SEQ ID NO:80, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:117, SEQ ID NO:135, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:148, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:170, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:180, SEQ ID NO:183, SEQ ID NO:192, SEQ ID NO:197, SEQ ID NO:218, SEQ ID NO:221, SEQ ID NO:230, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:251, SEQ ID NO:331, SEQ ID NO:334, SEQ ID NO:336, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:349, SEQ ID NO:358, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:373, SEQ ID NO:377, SEQ ID NO:380, SEQ ID NO:397, SEQ ID NO:399, SEQ ID NO:400, SEQ ID NO:401, SEQ ID NO:402, SEQ ID NO:403, SEQ ID NO:404, SEQ ID NO:405, SEQ ID NO:406, SEQ ID NO:407, SEQ ID NO:408, SEQ ID NO:409, SEQ ID NO:410, SEQ ID NO:411, SEQ ID NO:412, SEQ ID NO:413, SEQ ID NO:414, SEQ ID NO:415, SEQ ID NO:416, SEQ ID NO:417, SEQ ID NO:418, SEQ ID NO:419, SEQ ID NO:420, SEQ ID NO:421, SEQ ID NO:422, SEQ ID NO:423, SEQ ID NO:424, SEQ ID NO:425, SEQ ID NO:426, SEQ ID NO:427, SEQ ID NO:428, SEQ ID NO:429, SEQ ID NO:430, SEQ ID NO:431, SEQ ID NO:432, SEQ ID NO:433, SEQ ID NO:434, SEQ ID NO:435, SEQ ID NO:436, SEQ ID NO:437, SEQ ID NO:438, SEQ ID NO:439, SEQ ID NO:440, SEQ ID NO:441, SEQ ID NO:442, SEQ ID NO:443, SEQ ID NO:444, SEQ ID NO:445, SEQ ID NO:446, SEQ ID NO:447, SEQ ID NO:448, SEQ ID NO:449, SEQ ID NO:450, SEQ ID NO:451, SEQ ID NO:452, SEQ ID NO:453, SEQ ID NO:454, SEQ ID NO:455, SEQ ID NO:456, SEQ ID NO:457, SEQ ID NO:458, SEQ ID NO:459, SEQ ID NO:460, SEQ ID NO:461, SEQ ID NO:462, SEQ ID NO:463, SEQ ID NO:464, SEQ ID NO:465, SEQ ID NO:466, SEQ ID NO:467, SEQ ID NO:468, SEQ ID NO:469, SEQ ID NO:470, SEQ ID NO:471, SEQ ID NO:472, SEQ ID NO:473, SEQ ID NO:474, SEQ ID NO:475, SEQ ID NO:476, SEQ ID NO:477, SEQ ID NO:478, SEQ ID NO:479, SEQ ID NO:480, SEQ ID NO:481, SEQ ID NO:482, SEQ ID NO:483, SEQ ID NO:484, SEQ ID NO:485, SEQ ID NO:486, SEQ ID NO:487, SEQ ID NO:488, SEQ ID NO:489, SEQ ID NO:490, SEQ ID NO:491, SEQ ID NO:492, SEQ ID NO:493, SEQ ID NO:494, SEQ ID NO:495, SEQ ID NO:496, SEQ ID NO:497, SEQ ID NO:498, SEQ ID NO:499, SEQ ID NO:500, SEQ ID NO:501, SEQ ID NO:546, SEQ ID NO:547, SEQ ID NO:548, SEQ ID NO:549, SEQ ID NO:550, SEQ ID NO:551, SEQ ID NO:552, SEQ ID NO:553, SEQ ID NO:554, SEQ ID NO:555, SEQ ID NO:556, SEQ ID NO:557, SEQ ID NO:558, SEQ ID NO:559, SEQ ID NO:560, SEQ ID NO:561, SEQ ID NO:562, SEQ ID NO:563, SEQ ID NO:564, SEQ ID NO:565, SEQ ID NO:566, SEQ ID NO:567, SEQ ID NO:568, SEQ ID NO:569, SEQ ID

NO:570, SEQ ID NO:571, SEQ ID NO:572, SEQ ID NO:573, SEQ ID NO:574, or SEQ ID NO:575. Alternatively, a oil-modulating nucleic acid can be a variant of the nucleic acid having the nucleotide sequence set forth in SEQ ID NO:80, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:117, SEQ ID NO:135, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:148, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:170, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:180, SEQ ID NO:183, SEQ ID NO:192, SEQ ID NO:197, SEQ ID NO:218, SEQ ID NO:221, SEQ ID NO:230, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:251, SEQ ID NO:331, SEQ ID NO:334, SEQ ID NO:336, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:349, SEQ ID NO:358, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:373, SEQ ID NO:377, SEQ ID NO:380, SEQ ID NO:397, SEQ ID NO:399, SEQ ID NO:400, SEQ ID NO:401, SEQ ID NO:402, SEQ ID NO:403, SEQ ID NO:404, SEQ ID NO:405, SEQ ID NO:406, SEQ ID NO:407, SEQ ID NO:408, SEQ ID NO:409, SEQ ID NO:410, SEQ ID NO:411, SEQ ID NO:412, SEQ ID NO:413, SEQ ID NO:414, SEQ ID NO:415, SEQ ID NO:416, SEQ ID NO:417, SEQ ID NO:418, SEQ ID NO:419, SEQ ID NO:420, SEQ ID NO:421, SEQ ID NO:422, SEQ ID NO:423, SEQ ID NO:424, SEQ ID NO:425, SEQ ID NO:426, SEQ ID NO:427, SEQ ID NO:428, SEQ ID NO:429, SEQ ID NO:430, SEQ ID NO:431, SEQ ID NO:432, SEQ ID NO:433, SEQ ID NO:434, SEQ ID NO:435, SEQ ID NO:436, SEQ ID NO:437, SEQ ID NO:438, SEQ ID NO:439, SEQ ID NO:440, SEQ ID NO:441, SEQ ID NO:442, SEQ ID NO:443, SEQ ID NO:444, SEQ ID NO:445, SEQ ID NO:446, SEQ ID NO:447, SEQ ID NO:448, SEQ ID NO:449, SEQ ID NO:450, SEQ ID NO:451, SEQ ID NO:452, SEQ ID NO:453, SEQ ID NO:454, SEQ ID NO:455, SEQ ID NO:456, SEQ ID NO:457, SEQ ID NO:458, SEQ ID NO:459, SEQ ID NO:460, SEQ ID NO:461, SEQ ID NO:462, SEQ ID NO:463, SEQ ID NO:464, SEQ ID NO:465, SEQ ID NO:466, SEQ ID NO:467, SEQ ID NO:468, SEQ ID NO:469, SEQ ID NO:470, SEQ ID NO:471, SEQ ID NO:472, SEQ ID NO:473, SEQ ID NO:474, SEQ ID NO:475, SEQ ID NO:476, SEQ ID NO:477, SEQ ID NO:478, SEQ ID NO:479, SEQ ID NO:480, SEQ ID NO:481, SEQ ID NO:482, SEQ ID NO:483, SEQ ID NO:484, SEQ ID NO:485, SEQ ID NO:486, SEQ ID NO:487, SEQ ID NO:488, SEQ ID NO:489, SEQ ID NO:490, SEQ ID NO:491, SEQ ID NO:492, SEQ ID NO:493, SEQ ID NO:494, SEQ ID NO:495, SEQ ID NO:496, SEQ ID NO:497, SEQ ID NO:498, SEQ ID NO:499, SEQ ID NO:500, SEQ ID NO:501, SEQ ID NO:546, SEQ ID NO:547, SEQ ID NO:548, SEQ ID NO:549, SEQ ID NO:550, SEQ ID NO:551, SEQ ID NO:552, SEQ ID NO:553, SEQ ID NO:554, SEQ ID NO:555, SEQ ID NO:556, SEQ ID NO:557, SEQ ID NO:558, SEQ ID NO:559, SEQ ID NO:560, SEQ ID NO:561, SEQ ID NO:562, SEQ ID NO:563, SEQ ID NO:564, SEQ ID NO:565, SEQ ID NO:566, SEQ ID NO:567, SEQ ID NO:568, SEQ ID NO:569, SEQ ID NO:570, SEQ ID NO:571, SEQ ID NO:572, SEQ ID NO:573, SEQ ID NO:574, or SEQ ID NO:575. For example, a oil-modulating nucleic acid can have a nucleotide sequence with at least 80% sequence identity, e.g., 81%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity, to the nucleotide sequence set forth in SEQ ID NO:80, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:108,

SEQ ID NO:110, SEQ ID NO:117, SEQ ID NO:135, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:148, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:170, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:180, SEQ ID NO:183, SEQ ID NO:192, SEQ ID NO:197, SEQ ID NO:218, SEQ ID NO:221, SEQ ID NO:230, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:251, SEQ ID NO:331, SEQ ID NO:334, SEQ ID NO:336, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:349, SEQ ID NO:358, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:373, SEQ ID NO:377, SEQ ID NO:380, SEQ ID NO:397, SEQ ID NO:399, SEQ ID NO:400, SEQ ID NO:401, SEQ ID NO:402, SEQ ID NO:403, SEQ ID NO:404, SEQ ID NO:405, SEQ ID NO:406, SEQ ID NO:407, SEQ ID NO:408, SEQ ID NO:409, SEQ ID NO:410, SEQ ID NO:411, SEQ ID NO:412, SEQ ID NO:413, SEQ ID NO:414, SEQ ID NO:415, SEQ ID NO:416, SEQ ID NO:417, SEQ ID NO:418, SEQ ID NO:419, SEQ ID NO:420, SEQ ID NO:421, SEQ ID NO:422, SEQ ID NO:423, SEQ ID NO:424, SEQ ID NO:425, SEQ ID NO:426, SEQ ID NO:427, SEQ ID NO:428, SEQ ID NO:429, SEQ ID NO:430, SEQ ID NO:431, SEQ ID NO:432, SEQ ID NO:433, SEQ ID NO:434, SEQ ID NO:435, SEQ ID NO:436, SEQ ID NO:437, SEQ ID NO:438, SEQ ID NO:439, SEQ ID NO:440, SEQ ID NO:441, SEQ ID NO:442, SEQ ID NO:443, SEQ ID NO:444, SEQ ID NO:445, SEQ ID NO:446, SEQ ID NO:447, SEQ ID NO:448, SEQ ID NO:449, SEQ ID NO:450, SEQ ID NO:451, SEQ ID NO:452, SEQ ID NO:453, SEQ ID NO:454, SEQ ID NO:455, SEQ ID NO:456, SEQ ID NO:457, SEQ ID NO:458, SEQ ID NO:459, SEQ ID NO:460, SEQ ID NO:461, SEQ ID NO:462, SEQ ID NO:463, SEQ ID NO:464, SEQ ID NO:465, SEQ ID NO:466, SEQ ID NO:467, SEQ ID NO:468, SEQ ID NO:469, SEQ ID NO:470, SEQ ID NO:471, SEQ ID NO:472, SEQ ID NO:473, SEQ ID NO:474, SEQ ID NO:475, SEQ ID NO:476, SEQ ID NO:477, SEQ ID NO:478, SEQ ID NO:479, SEQ ID NO:480, SEQ ID NO:481, SEQ ID NO:482, SEQ ID NO:483, SEQ ID NO:484, SEQ ID NO:485, SEQ ID NO:486, SEQ ID NO:487, SEQ ID NO:488, SEQ ID NO:489, SEQ ID NO:490, SEQ ID NO:491, SEQ ID NO:492, SEQ ID NO:493, SEQ ID NO:494, SEQ ID NO:495, SEQ ID NO:496, SEQ ID NO:497, SEQ ID NO:498, SEQ ID NO:499, SEQ ID NO:500, SEQ ID NO:501, SEQ ID NO:546, SEQ ID NO:547, SEQ ID NO:548, SEQ ID NO:549, SEQ ID NO:550, SEQ ID NO:551, SEQ ID NO:552, SEQ ID NO:553, SEQ ID NO:554, SEQ ID NO:555, SEQ ID NO:556, SEQ ID NO:557, SEQ ID NO:558, SEQ ID NO:559, SEQ ID NO:560, SEQ ID NO:561, SEQ ID NO:562, SEQ ID NO:563, SEQ ID NO:564, SEQ ID NO:565, SEQ ID NO:566, SEQ ID NO:567, SEQ ID NO:568, SEQ ID NO:569, SEQ ID NO:570, SEQ ID NO:571, SEQ ID NO:572, SEQ ID NO:573, SEQ ID NO:574, or SEQ ID NO:575.

An "isolated nucleic acid" can be, for example, a naturally-occurring DNA molecule, provided one of the nucleic acid sequences normally found immediately flanking that DNA molecule in a naturally-occurring genome is removed or absent. Thus, an isolated nucleic acid includes, without limitation, a DNA molecule that exists as a separate molecule, independent of other sequences (e.g., a chemically synthesized nucleic acid, or a cDNA or genomic DNA fragment produced by the polymerase chain reaction (PCR) or restric-



tion endonuclease treatment). An isolated nucleic acid also refers to a DNA molecule that is incorporated into a vector, an autonomously replicating plasmid, a virus, or into the genomic DNA of a prokaryote or eukaryote. In addition, an isolated nucleic acid can include an engineered nucleic acid such as a DNA molecule that is part of a hybrid or fusion nucleic acid. A nucleic acid existing among hundreds to millions of other nucleic acids within, for example, cDNA libraries or genomic libraries, or gel slices containing a genomic DNA restriction digest, is not to be considered an isolated nucleic acid.

Isolated nucleic acid molecules can be produced by standard techniques. For example, polymerase chain reaction (PCR) techniques can be used to obtain an isolated nucleic acid containing a nucleotide sequence described herein. PCR can be used to amplify specific sequences from DNA as well as RNA, including sequences from total genomic DNA or total cellular RNA. Various PCR methods are described, for example, in *PCR Primer: A Laboratory Manual*, Dieffenbach and Dveksler, eds., Cold Spring Harbor Laboratory Press, 1995. Generally, sequence information from the ends of the region of interest or beyond is employed to design oligonucleotide primers that are identical or similar in sequence to opposite strands of the template to be amplified. Various PCR strategies also are available by which site-specific nucleotide sequence modifications can be introduced into a template nucleic acid. Isolated nucleic acids also can be chemically synthesized, either as a single nucleic acid molecule (e.g., using automated DNA synthesis in the 3' to 5' direction using phosphoramidite technology) or as a series of oligonucleotides. For example, one or more pairs of long oligonucleotides (e.g., >100 nucleotides) can be synthesized that contain the desired sequence, with each pair containing a short segment of complementarity (e.g., about 15 nucleotides) such that a duplex is formed when the oligonucleotide pair is annealed. DNA polymerase is used to extend the oligonucleotides, resulting in a single, double-stranded nucleic acid molecule per oligonucleotide pair, which then can be ligated into a vector. Isolated nucleic acids of the invention also can be obtained by mutagenesis of, e.g., a naturally occurring DNA.

As used herein, the term “percent sequence identity” refers to the degree of identity between any given query sequence, e.g., SEQ ID NO:81, and a subject sequence. A subject sequence typically has a length that is from 80 percent to 200 percent of the length of the query sequence, e.g., 82, 85, 87, 89, 90, 93, 95, 97, 99, 100, 105, 110, 115, 120, 130, 140, 150, 160, 170, 180, 190, or 200 percent of the length of the query sequence. A percent identity for any subject nucleic acid or polypeptide relative to a query nucleic acid or polypeptide can be determined as follows. A query sequence (e.g., a nucleic acid sequence or an amino acid sequence) is aligned to one or more subject sequences using the computer program ClustalW (version 1.83, default parameters), which allows alignments of nucleic acid or polypeptide sequences to be carried out across their entire length (global alignment). Chenna et al., *Nucleic Acids Res.*, 31(13):3497-500 (2003).

ClustalW calculates the best match between a query and one or more subject sequences, and aligns them so that identities, similarities and differences can be determined. Gaps of one or more residues can be inserted into a query sequence, a subject sequence, or both, to maximize sequence alignments. For fast pairwise alignment of nucleic acid sequences, the following default parameters are used: word size: 2; window size: 4; scoring method: percentage; number of top diagonals: 4; and gap penalty: 5. For multiple alignment of nucleic acid sequences, the following parameters are used: gap opening

penalty: 10.0; gap extension penalty: 5.0; and weight transitions: yes. For fast pairwise alignment of protein sequences, the following parameters are used: word size: 1; window size: 5; scoring method: percentage; number of top diagonals: 5; gap penalty: 3. For multiple alignment of protein sequences, the following parameters are used: weight matrix: blosum; gap opening penalty: 10.0; gap extension penalty: 0.05; hydrophilic gaps: on; hydrophilic residues: Gly, Pro, Ser, Asn, Asp, Gln, Glu, Arg, and Lys; residue-specific gap penalties: on. The ClustalW output is a sequence alignment that reflects the relationship between sequences. ClustalW can be run, for example, at the Baylor College of Medicine Search Launcher site ([searchlauncher.bcm.tmc.edu/multi-align/multi-align.html](http://searchlauncher.bcm.tmc.edu/multi-align/multi-align.html)) and at the European Bioinformatics Institute site on the World Wide Web ([ebi.ac.uk/clustalw](http://ebi.ac.uk/clustalw)).

To determine percent identity of a subject nucleic acid or amino acid sequence to a query sequence, the sequences are aligned using ClustalW, the number of identical matches in the alignment is divided by the length of the query sequence, and the result is multiplied by 100. It is noted that the percent identity value can be rounded to the nearest tenth. For example, 78.11, 78.12, 78.13, and 78.14 are rounded down to 78.1, while 78.15, 78.16, 78.17, 78.18, and 78.19 are rounded up to 78.2.

The term “exogenous” with respect to a nucleic acid indicates that the nucleic acid is part of a recombinant nucleic acid construct, or is not in its natural environment. For example, an exogenous nucleic acid can be a sequence from one species introduced into another species, i.e., a heterologous nucleic acid. Typically, such an exogenous nucleic acid is introduced into the other species via a recombinant nucleic acid construct. An exogenous nucleic acid can also be a sequence that is native to an organism and that has been reintroduced into cells of that organism. An exogenous nucleic acid that includes a native sequence can often be distinguished from the naturally occurring sequence by the presence of non-natural sequences linked to the exogenous nucleic acid, e.g., non-native regulatory sequences flanking a native sequence in a recombinant nucleic acid construct. In addition, stably transformed exogenous nucleic acids typically are integrated at positions other than the position where the native sequence is found. It will be appreciated that an exogenous nucleic acid may have been introduced into a progenitor and not into the cell under consideration. For example, a transgenic plant containing an exogenous nucleic acid can be the progeny of a cross between a stably transformed plant and a non-transgenic plant. Such progeny are considered to contain the exogenous nucleic acid.

Recombinant constructs are also provided herein and can be used to transform plants or plant cells in order to modulate oil levels. A recombinant nucleic acid construct comprises a nucleic acid encoding an oil-modulating polypeptide as described herein, operably linked to a regulatory region suitable for expressing the oil-modulating polypeptide in the plant or cell. Thus, a nucleic acid can comprise a coding sequence that encodes any of the oil-modulating polypeptides as set forth in SEQ ID NOs:81-92, SEQ ID NO:94, SEQ ID NOs:96-107, SEQ ID NO:109, SEQ ID NOs:111-116, SEQ ID NOs:118-134, SEQ ID NOs:136-140, SEQ ID NO:142, SEQ ID NO:144, SEQ ID NOs:146-147, SEQ ID NOs:149-150, SEQ ID NO:152, SEQ ID NO:154, SEQ ID NOs:156-157, SEQ ID NOs:159-169, SEQ ID NOs:171-172, SEQ ID NO:174, SEQ ID NO:176, SEQ ID NOs:178-179, SEQ ID NOs:181-182, SEQ ID NOs:184-191, SEQ ID NOs:193-196, SEQ ID NOs:198-217, SEQ ID NOs:219-220, SEQ ID NOs:222-229, SEQ ID NOs:231-238, SEQ ID NO:240, SEQ ID NOs:242-250, SEQ ID NOs:252-330, SEQ ID NOs:332-333,

SEQ ID NO:335, SEQ ID NOs:337-340, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346-348, SEQ ID NOs:350-357, SEQ ID NOs:359-366, SEQ ID NO:368, SEQ ID NOs:370-372, SEQ ID NOs:374-376, SEQ ID NOs:378-379, SEQ ID NOs:381-396, SEQ ID NO:398, SEQ ID NOs:502-545, and the consensus sequences set forth in FIGS. 1-13. Examples of nucleic acids encoding oil-modulating polypeptides are set forth in SEQ ID NO:80, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:108, SEQ ID NO:110, SEQ ID NO:117, SEQ ID NO:135, SEQ ID NO:141, SEQ ID NO:143, SEQ ID NO:145, SEQ ID NO:148, SEQ ID NO:151, SEQ ID NO:153, SEQ ID NO:155, SEQ ID NO:158, SEQ ID NO:170, SEQ ID NO:173, SEQ ID NO:175, SEQ ID NO:177, SEQ ID NO:180, SEQ ID NO:183, SEQ ID NO:192, SEQ ID NO:197, SEQ ID NO:218, SEQ ID NO:221, SEQ ID NO:230, SEQ ID NO:239, SEQ ID NO:241, SEQ ID NO:251, SEQ ID NO:331, SEQ ID NO:334, SEQ ID NO:336, SEQ ID NO:341, SEQ ID NO:343, SEQ ID NO:345, SEQ ID NO:349, SEQ ID NO:358, SEQ ID NO:367, SEQ ID NO:369, SEQ ID NO:373, SEQ ID NO:377, SEQ ID NO:380, SEQ ID NO:397, SEQ ID NOs:399-501, and SEQ ID NOs:546-575.

In some cases, a recombinant nucleic acid construct can include a nucleic acid comprising less than the full-length of a coding sequence. Typically, such a construct also includes a regulatory region operably linked to the oil-modulating nucleic acid. In some cases, a recombinant nucleic acid construct can include a nucleic acid comprising a coding sequence, a gene, or a fragment of a coding sequence or gene in an antisense orientation so that the antisense strand of RNA is transcribed.

It will be appreciated that a number of nucleic acids can encode a polypeptide having a particular amino acid sequence. The degeneracy of the genetic code is well known to the art; i.e., for many amino acids, there is more than one nucleotide triplet that serves as the codon for the amino acid. For example, codons in the coding sequence for a given oil-modulating polypeptide can be modified such that optimal expression in a particular plant species is obtained, using appropriate codon bias tables for that species.

Vectors containing nucleic acids such as those described herein also are provided. A "vector" is a replicon, such as a plasmid, phage, or cosmid, into which another DNA segment may be inserted so as to bring about the replication of the inserted segment. Generally, a vector is capable of replication when associated with the proper control elements. Suitable vector backbones include, for example, those routinely used in the art such as plasmids, viruses, artificial chromosomes, BACs, YACs, or PACs. The term "vector" includes cloning and expression vectors, as well as viral vectors and integrating vectors. An "expression vector" is a vector that includes a regulatory region. Suitable expression vectors include, without limitation, plasmids and viral vectors derived from, for example, bacteriophage, baculoviruses, and retroviruses. Numerous vectors and expression systems are commercially available from such corporations as Novagen (Madison, Wis.), Clontech (Palo Alto, Calif.), Stratagene (La Jolla, Calif.), and Invitrogen/Life Technologies (Carlsbad, Calif.).

The vectors provided herein also can include, for example, origins of replication, scaffold attachment regions (SARs), and/or markers. A marker gene can confer a selectable phenotype on a plant cell. For example, a marker can confer biocide resistance, such as resistance to an antibiotic (e.g., kanamycin, G418, bleomycin, or hygromycin), or an herbicide (e.g., chlorosulfuron or phosphinothricin). In addition, an expression vector can include a tag sequence designed to facilitate manipulation or detection (e.g., purification or

localization) of the expressed polypeptide. Tag sequences, such as green fluorescent protein (GFP), glutathione S-transferase (GST), polyhistidine, c-myc, hemagglutinin, or Flag™ tag (Kodak, New Haven, Conn.) sequences typically are expressed as a fusion with the encoded polypeptide. Such tags can be inserted anywhere within the polypeptide, including at either the carboxyl or amino terminus.

#### Regulatory Regions

The term "regulatory region" refers to nucleotide sequences that influence transcription or translation initiation and rate, and stability and/or mobility of a transcription or translation product. Regulatory regions include, without limitation, promoter sequences, enhancer sequences, response elements, protein recognition sites, inducible elements, protein binding sequences, 5' and 3' untranslated regions (UTRs), transcriptional start sites, termination sequences, polyadenylation sequences, introns, and combinations thereof.

As used herein, the term "operably linked" refers to positioning of a regulatory region and a sequence to be transcribed in a nucleic acid so as to influence transcription or translation of such a sequence. For example, to bring a coding sequence under the control of a promoter, the translation initiation site of the translational reading frame of the polypeptide is typically positioned between one and about fifty nucleotides downstream of the promoter. A promoter can, however, be positioned as much as about 5,000 nucleotides upstream of the translation initiation site, or about 2,000 nucleotides upstream of the transcription start site. A promoter typically comprises at least a core (basal) promoter. A promoter also may include at least one control element, such as an enhancer sequence, an upstream element or an upstream activation region (UAR). For example, a suitable enhancer is a cis-regulatory element (-212 to -154) from the upstream region of the octopine synthase (ocs) gene. Fromm et al., *The Plant Cell*, 1:977-984 (1989). The choice of promoters to be included depends upon several factors, including, but not limited to, efficiency, selectability, inducibility, desired expression level, and cell- or tissue-preferential expression. It is a routine matter for one of skill in the art to modulate the expression of a coding sequence by appropriately selecting and positioning regulatory regions relative to the coding sequence.

Some suitable regulatory regions initiate transcription only, or predominantly, in certain cell types, for example, a promoter that is active predominantly in a reproductive tissue such as fruit, ovule, pollen, pistils, female gametophyte, egg cell, central cell, nucellus, suspensor, synergid cell, flowers, embryonic tissue, embryo sac, embryo, zygote, endosperm, integument, or seed coat. Thus, as used herein a cell type- or tissue-preferential promoter is one that drives expression preferentially in the target tissue, but may also lead to some expression in other cell types or tissues as well. Methods for identifying and characterizing promoter regions in plant genomic DNA include, for example, those described in the following references: Jordano et al., *Plant Cell*, 1:855-866 (1989); Bustos et al., *Plant Cell*, 1:839-854 (1989); Green et al., *EMBO J.*, 7:4035-4044 (1988); Meier et al., *Plant Cell*, 3:309-316 (1991); and Zhang et al., *Plant Physiology*, 110:1069-1079 (1996).

Examples of various classes of regulatory regions are described below. Some of the regulatory regions indicated below as well as additional regulatory regions are described in more detail in U.S. Patent Application Ser. Nos. 60/505,689; 60/518,075; 60/544,771; 60/558,869; 60/583,691; 60/619,181; 60/637,140; 60/757,544; 60/776,307; Ser. Nos. 10/957,569; 11/058,689; 11/172,703; 11/208,308; 11/274,890; 60/583,609; 60/612,891; 11/097,589; 11/233,726;

11/408,791; 11/414,142; 10/950,321; 11/360,017; PCT/US05/011105; PCT/US05/034308; and PCT/US05/23639. Nucleotide sequences of promoters are set forth in SEQ ID NOs:1-79 and 259-274. It will be appreciated that a regulatory region may meet criteria for one classification based on its activity in one plant species, and yet meet criteria for a different classification based on its activity in another plant species.

#### Broadly Expressing Promoters

A promoter can be said to be "broadly expressing" when it promotes transcription in many, but not necessarily all, plant tissues. For example, a broadly expressing promoter can promote transcription of an operably linked sequence in one or more of the shoot, shoot tip (apex), and leaves, but weakly or not at all in tissues such as roots or stems. As another example, a broadly expressing promoter can promote transcription of an operably linked sequence in one or more of the stem, shoot, shoot tip (apex), and leaves, but can promote transcription weakly or not at all in tissues such as reproductive tissues of flowers and developing seeds. Non-limiting examples of broadly expressing promoters that can be included in the nucleic acid constructs provided herein include the p326 (SEQ ID NO:76), YP0144 (SEQ ID NO:55), YP0190 (SEQ ID NO:59), p13879 (SEQ ID NO:75), YP0050 (SEQ ID NO:35), p32449 (SEQ ID NO:77), 21876 (SEQ ID NO:1), YP0158 (SEQ ID NO:57), YP0214 (SEQ ID NO:61), YP0380 (SEQ ID NO:70), PT0848 (SEQ ID NO:26), and PT0633 (SEQ ID NO:7) promoters. Additional examples include the cauliflower mosaic virus (CaMV) 35S promoter, the mannopine synthase (MAS) promoter, the 1' or 2' promoters derived from T-DNA of *Agrobacterium tumefaciens*, the figwort mosaic virus 34S promoter, actin promoters such as the rice actin promoter, and ubiquitin promoters such as the maize ubiquitin-1 promoter. In some cases, the CaMV 35S promoter is excluded from the category of broadly expressing promoters.

#### Root Promoters

Root-active promoters confer transcription in root tissue, e.g., root endodermis, root epidermis, or root vascular tissues. In some embodiments, root-active promoters are root-preferential promoters, i.e., confer transcription only or predominantly in root tissue. Root-preferential promoters include the YP0128 (SEQ ID NO:52), YP0275 (SEQ ID NO:63), PT0625 (SEQ ID NO:6), PT0660 (SEQ ID NO:9), PT0683 (SEQ ID NO:14), and PT0758 (SEQ ID NO:22) promoters. Other root-preferential promoters include the PT0613 (SEQ ID NO:5), PT0672 (SEQ ID NO:11), PT0688 (SEQ ID NO:15), and PT0837 (SEQ ID NO:24) promoters, which drive transcription primarily in root tissue and to a lesser extent in ovules and/or seeds. Other examples of root-preferential promoters include the root-specific subdomains of the CaMV 35S promoter (Lam et al., *Proc. Natl. Acad. Sci. USA*, 86:7890-7894 (1989)), root cell specific promoters reported by Conkling et al., *Plant Physiol.*, 93:1203-1211 (1990), and the tobacco RD2 promoter.

#### Maturing Endosperm Promoters

In some embodiments, promoters that drive transcription in maturing endosperm can be useful. Transcription from a maturing endosperm promoter typically begins after fertilization and occurs primarily in endosperm tissue during seed development and is typically highest during the cellularization phase. Most suitable are promoters that are active predominantly in maturing endosperm, although promoters that are also active in other tissues can sometimes be used. Non-limiting examples of maturing endosperm promoters that can be included in the nucleic acid constructs provided herein include the napin promoter, the Arcelin-5 promoter, the

phaseolin promoter (Bustos et al., *Plant Cell*, 1(9):839-853 (1989)), the soybean trypsin inhibitor promoter (Riggs et al., *Plant Cell*, 1(6):609-621 (1989)), the ACP promoter (Baerson et al., *Plant Mol. Biol.*, 22(2):255-267 (1993)), the stearyl-ACP desaturase promoter (Slocombe et al., *Plant Physiol.*, 104(4):167-176 (1994)), the soybean  $\alpha'$  subunit of  $\beta$ -conglycinin promoter (Chen et al., *Proc. Natl. Acad. Sci. USA*, 83:8560-8564 (1986)), the oleosin promoter (Hong et al., *Plant Mol. Biol.*, 34(3):549-555 (1997)), and zein promoters, such as the 15 kD zein promoter, the 16 kD zein promoter, 19 kD zein promoter, 22 kD zein promoter and 27 kD zein promoter. Also suitable are the Osgt-1 promoter from the rice glutelin-1 gene (Zheng et al., *Mol. Cell Biol.*, 13:5829-5842 (1993)), the beta-amylase promoter, and the barley hordein promoter. Other maturing endosperm promoters include the YP0092 (SEQ ID NO:38), PT0676 (SEQ ID NO:12), and PT0708 (SEQ ID NO:17) promoters.

#### Ovary Tissue Promoters

Promoters that are active in ovary tissues such as the ovule wall and mesocarp can also be useful, e.g., a polygalacturonidase promoter, the banana TRX promoter, the melon actin promoter, YP0396 (SEQ ID NO:74), and PT0623 (SEQ ID NO:273). Examples of promoters that are active primarily in ovules include YP0007 (SEQ ID NO:30), YP0111 (SEQ ID NO:46), YP0092 (SEQ ID NO:38), YP0103 (SEQ ID NO:43), YP0028 (SEQ ID NO:33), YP0121 (SEQ ID NO:51), YP0008 (SEQ ID NO:31), YP0039 (SEQ ID NO:34), YP0115 (SEQ ID NO:47), YP0119 (SEQ ID NO:49), YP0120 (SEQ ID NO:50), and YP0374 (SEQ ID NO:68).

#### Embryo Sac/Early Endosperm Promoters

To achieve expression in embryo sac/early endosperm, regulatory regions can be used that are active in polar nuclei and/or the central cell, or in precursors to polar nuclei, but not in egg cells or precursors to egg cells. Most suitable are promoters that drive expression only or predominantly in polar nuclei or precursors thereto and/or the central cell. A pattern of transcription that extends from polar nuclei into early endosperm development can also be found with embryo sac/early endosperm-preferential promoters, although transcription typically decreases significantly in later endosperm development during and after the cellularization phase. Expression in the zygote or developing embryo typically is not present with embryo sac/early endosperm promoters.

Promoters that may be suitable include those derived from the following genes: *Arabidopsis* viviparous-1 (see, GenBank No. U93215); *Arabidopsis* atmyc1 (see, Urao (1996) *Plant Mol. Biol.*, 32:571-57; Conceicao (1994) *Plant*, 5:493-505); *Arabidopsis* FIE (GenBank No. AF129516); *Arabidopsis* MEA; *Arabidopsis* FIS2 (GenBank No. AF096096); and FIE 1.1 (U.S. Pat. No. 6,906,244). Other promoters that may be suitable include those derived from the following genes: maize MAC1 (see, Sheridan (1996) *Genetics*, 142:1009-1020); maize Cat3 (see, GenBank No. L05934; Abler (1993) *Plant Mol. Biol.*, 22:10131-1038). Other promoters include the following *Arabidopsis* promoters: YP0039 (SEQ ID NO:34), YP0101 (SEQ ID NO:41), YP0102 (SEQ ID NO:42), YP0110 (SEQ ID NO:45), YP0117 (SEQ ID NO:48), YP0119 (SEQ ID NO:49), YP0137 (SEQ ID NO:53), DME, YP0285 (SEQ ID NO:64), and YP0212 (SEQ ID NO:60). Other promoters that may be useful include the following rice promoters: p530c10 (SEQ ID NO:576), pOsFIE2-2 (SEQ ID NO:577), pOsMEA (SEQ ID NO:578), pOsYp102 (SEQ ID NO:579), and pOsYp285 (SEQ ID NO:580).

## Embryo Promoters

Regulatory regions that preferentially drive transcription in zygotic cells following fertilization can provide embryo-preferential expression. Most suitable are promoters that preferentially drive transcription in early stage embryos prior to the heart stage, but expression in late stage and maturing embryos is also suitable. Embryo-preferential promoters include the barley lipid transfer protein (Ltp1) promoter (*Plant Cell Rep.* (2001) 20:647-654), YP0097 (SEQ ID NO:40), YP0107 (SEQ ID NO:44), YP0088 (SEQ ID NO:37), YP0143 (SEQ ID NO:54), YP0156 (SEQ ID NO:56), PT0650 (SEQ ID NO:8), PT0695 (SEQ ID NO:16), PT0723 (SEQ ID NO:19), PT0838 (SEQ ID NO:25), PT0879 (SEQ ID NO:28), and PT0740 (SEQ ID NO:20).

## Photosynthetic Tissue Promoters

Promoters active in photosynthetic tissue confer transcription in green tissues such as leaves and stems. Most suitable are promoters that drive expression only or predominantly in such tissues. Examples of such promoters include the ribulose-1,5-bisphosphate carboxylase (RbcS) promoters such as the RbcS promoter from eastern larch (*Larix laricina*), the pine cab6 promoter (Yamamoto et al., *Plant Cell Physiol.*, 35:773-778 (1994)), the Cab-1 promoter from wheat (Fejes et al., *Plant Mol. Biol.*, 15:921-932 (1990)), the CAB-1 promoter from spinach (Lubberstedt et al., *Plant Physiol.*, 104:997-1006 (1994)), the cab IR promoter from rice (Luan et al., *Plant Cell*, 4:971-981 (1992)), the pyruvate orthophosphate dikinase (PPDK) promoter from corn (Matsuoka et al., *Proc. Natl. Acad. Sci. USA*, 90:9586-9590 (1993)), the tobacco Lhcb1\*2 promoter (Cerdan et al., *Plant Mol. Biol.*, 33:245-255 (1997)), the *Arabidopsis thaliana* SUC2 sucrose-H<sup>+</sup> symporter promoter (Truernit et al., *Planta*, 196:564-570 (1995)), and thylakoid membrane protein promoters from spinach (psaD, psaF, psaE, PC, FNR, atpC, atpD, cab, rbcS). Other photosynthetic tissue promoters include PT0535 (SEQ ID NO:3), PT0668 (SEQ ID NO:2), PT0886 (SEQ ID NO:29), PR0924 (SEQ ID NO:78), YP0144 (SEQ ID NO:55), YP0380 (SEQ ID NO:70), and PT0585 (SEQ ID NO:4).

## Vascular Tissue Promoters

Examples of promoters that have high or preferential activity in vascular bundles include YP0087 (SEQ ID NO:583), YP0093 (SEQ ID NO:584), YP0108 (SEQ ID NO:585), YP0022 (SEQ ID NO:586), and YP0080 (SEQ ID NO:587). Other vascular tissue-preferential promoters include the glycine-rich cell wall protein GRP 1.8 promoter (Keller and Baumgartner, *Plant Cell*, 3(10):1051-1061 (1991)), the *Commelina yellow mottle virus* (COYMV) promoter (Medberry et al., *Plant Cell*, 4(2):185-192 (1992)), and the rice tungro bacilliform virus (RTBV) promoter (Dai et al., *Proc. Natl. Acad. Sci. USA*, 101(2):687-692 (2004)).

## Inducible Promoters

Inducible promoters confer transcription in response to external stimuli such as chemical agents or environmental stimuli. For example, inducible promoters can confer transcription in response to hormones such as gibberellic acid or ethylene, or in response to light or drought. Examples of drought-inducible promoters include YP0380 (SEQ ID NO:70), PT0848 (SEQ ID NO:26), YP0381 (SEQ ID NO:71), YP0337 (SEQ ID NO:66), PT0633 (SEQ ID NO:7), YP0374 (SEQ ID NO:68), PT0710 (SEQ ID NO:18), YP0356 (SEQ ID NO:67), YP0385 (SEQ ID NO:73), YP0396 (SEQ ID NO:74), YP0388 (SEQ ID NO:588), YP0384 (SEQ ID NO:72), PT0688 (SEQ ID NO:15), YP0286 (SEQ ID NO:65), YP0377 (SEQ ID NO:69), PD1367 (SEQ ID NO:79), and PD0901 (SEQ ID NO:589). Nitrogen-inducible promoters include PT0863 (SEQ ID NO:27), PT0829 (SEQ

ID NO:23), PT0665 (SEQ ID NO:10), and PT0886 (SEQ ID NO:29). Examples of a shade-inducible promoters are PR0924 (SEQ ID NO:78) and PT0678 (SEQ ID NO:13).

## Basal Promoters

A basal promoter is the minimal sequence necessary for assembly of a transcription complex required for transcription initiation. Basal promoters frequently include a "TATA box" element that may be located between about 15 and about 35 nucleotides upstream from the site of transcription initiation. Basal promoters also may include a "CCAAT box" element (typically the sequence CCAAT) and/or a GGGCG sequence, which can be located between about 40 and about 200 nucleotides, typically about 60 to about 120 nucleotides, upstream from the transcription start site.

## Other Promoters

Other classes of promoters include, but are not limited to, leaf-preferential, stem/shoot-preferential, callus-preferential, guard cell-preferential, such as PT0678 (SEQ ID NO:13), tuber-preferential, parenchyma cell-preferential, and senescence-preferential promoters. Promoters designated YP0086 (SEQ ID NO:36), YP0188 (SEQ ID NO:58), YP0263 (SEQ ID NO:62), PT0758 (SEQ ID NO:22), PT0743 (SEQ ID NO:21), PT0829 (SEQ ID NO:23), YP0119 (SEQ ID NO:49), and YP0096 (SEQ ID NO:39), as described in the above-referenced patent applications, may also be useful.

## Other Regulatory Regions

A 5' untranslated region (UTR) can be included in nucleic acid constructs described herein. A 5' UTR is transcribed, but is not translated, and lies between the start site of the transcript and the translation initiation codon and may include the +1 nucleotide. A 3' UTR can be positioned between the translation termination codon and the end of the transcript. UTRs can have particular functions such as increasing mRNA stability or attenuating translation. Examples of 3' UTRs include, but are not limited to, polyadenylation signals and transcription termination sequences, e.g., a nopaline synthase termination sequence.

It will be understood that more than one regulatory region may be present in a recombinant polynucleotide, e.g., introns, enhancers, upstream activation regions, transcription terminators, and inducible elements. Thus, for example, more than one regulatory region can be operably linked to the sequence of a polynucleotide encoding an oil-modulating polypeptide.

Regulatory regions, such as promoters for endogenous genes, can be obtained by chemical synthesis or by subcloning from a genomic DNA that includes such a regulatory region. A nucleic acid comprising such a regulatory region can also include flanking sequences that contain restriction enzyme sites that facilitate subsequent manipulation.

## Transgenic Plants and Plant Cells

The invention also features transgenic plant cells and plants comprising at least one recombinant nucleic acid construct described herein. A plant or plant cell can be transformed by having a construct integrated into its genome, i.e., can be stably transformed. Stably transformed cells typically retain the introduced nucleic acid with each cell division. A plant or plant cell can also be transiently transformed such that the construct is not integrated into its genome. Transiently transformed cells typically lose all or some portion of the introduced nucleic acid construct with each cell division such that the introduced nucleic acid cannot be detected in daughter cells after a sufficient number of cell divisions. Both transiently transformed and stably transformed transgenic plants and plant cells can be useful in the methods described herein.

Transgenic plant cells used in methods described herein can constitute part or all of a whole plant. Such plants can be

grown in a manner suitable for the species under consideration, either in a growth chamber, a greenhouse, or in a field. Transgenic plants can be bred as desired for a particular purpose, e.g., to introduce a recombinant nucleic acid into other lines, to transfer a recombinant nucleic acid to other species, or for further selection of other desirable traits. Alternatively, transgenic plants can be propagated vegetatively for those species amenable to such techniques. As used herein, a transgenic plant also refers to progeny of an initial transgenic plant. Progeny includes descendants of a particular plant or plant line. Progeny of an instant plant include seeds formed on  $F_1$ ,  $F_2$ ,  $F_3$ ,  $F_4$ ,  $F_5$ ,  $F_6$  and subsequent generation plants, or seeds formed on  $BC_1$ ,  $BC_2$ ,  $BC_3$ , and subsequent generation plants, or seeds formed on  $F_1BC_1$ ,  $F_1BC_2$ ,  $F_1BC_3$ , and subsequent generation plants. The designation  $F_1$  refers to the progeny of a cross between two parents that are genetically distinct. The designations  $F_2$ ,  $F_3$ ,  $F_4$ ,  $F_5$  and  $F_6$  refer to subsequent generations of self- or sib-pollinated progeny of an  $F_1$  plant. Seeds produced by a transgenic plant can be grown and then selfed (or outcrossed and selfed) to obtain seeds homozygous for the nucleic acid construct.

Transgenic plants can be grown in suspension culture, or tissue or organ culture. For the purposes of this invention, solid and/or liquid tissue culture techniques can be used. When using solid medium, transgenic plant cells can be placed directly onto the medium or can be placed onto a filter that is then placed in contact with the medium. When using liquid medium, transgenic plant cells can be placed onto a flotation device, e.g., a porous membrane that contacts the liquid medium. Solid medium typically is made from liquid medium by adding agar. For example, a solid medium can be Murashige and Skoog (MS) medium containing agar and a suitable concentration of an auxin, e.g., 2,4-dichlorophenoxyacetic acid (2,4-D), and a suitable concentration of a cytokinin, e.g., kinetin.

When transiently transformed plant cells are used, a reporter sequence encoding a reporter polypeptide having a reporter activity can be included in the transformation procedure and an assay for reporter activity or expression can be performed at a suitable time after transformation. A suitable time for conducting the assay typically is about 1-21 days after transformation, e.g., about 1-14 days, about 1-7 days, or about 1-3 days. The use of transient assays is particularly convenient for rapid analysis in different species, or to confirm expression of a heterologous oil-modulating polypeptide whose expression has not previously been confirmed in particular recipient cells.

Techniques for introducing nucleic acids into monocotyledonous and dicotyledonous plants are known in the art, and include, without limitation, *Agrobacterium*-mediated transformation, viral vector-mediated transformation, electroporation and particle gun transformation, e.g., U.S. Pat. Nos. 5,538,880; 5,204,253; 6,329,571 and 6,013,863. If a cell or cultured tissue is used as the recipient tissue for transformation, plants can be regenerated from transformed cultures if desired, by techniques known to those skilled in the art.

In aspects related to making transgenic plants, a typical step involves selection or screening of transformed plants, e.g., for the presence of a functional vector as evidenced by expression of a selectable marker. Selection or screening can be carried out among a population of recipient cells to identify transformants using selectable marker genes such as herbicide resistance genes. Physical and biochemical methods can be used to identify transformants. These include Southern analysis or PCR amplification for detection of a polynucleotide; Northern blots, S1 RNase protection, primer-extension, or RT-PCR amplification for detecting RNA transcripts;

enzymatic assays for detecting enzyme or ribozyme activity of polypeptides and polynucleotides; and protein gel electrophoresis, Western blots, immunoprecipitation, and enzyme-linked immunoassays to detect polypeptides. Other techniques such as in situ hybridization, enzyme staining, and immunostaining also can be used to detect the presence or expression of polypeptides and/or polynucleotides. Methods for performing all of the referenced techniques are known.

A population of transgenic plants can be screened and/or selected for those members of the population that have a desired trait or phenotype conferred by expression of the transgene. For example, a population of progeny of a single transformation event can be screened for those plants having a desired level of expression of a heterologous oil-modulating polypeptide or nucleic acid. As an alternative, a population of plants comprising independent transformation events can be screened for those plants having a desired trait, such as a modulated level of oil. Selection and/or screening can be carried out over one or more generations, which can be useful to identify those plants that have a statistically significant difference in an oil level as compared to a corresponding level in a control plant. Selection and/or screening can also be carried out in more than one geographic location. In some cases, transgenic plants can be grown and selected under conditions which induce a desired phenotype or are otherwise necessary to produce a desired phenotype in a transgenic plant. In addition, selection and/or screening can be carried out during a particular developmental stage in which the phenotype is expected to be exhibited by the plant. Selection and/or screening can be carried out to choose those transgenic plants having a statistically significant difference in an oil level relative to a control plant that lacks the transgene. Selected or screened transgenic plants have an altered phenotype as compared to a corresponding control plant, as described in the "Transgenic Plant Phenotypes" section below.

#### Plant Species

The polynucleotides and vectors described herein can be used to transform a number of monocotyledonous and dicotyledonous plants and plant cell systems, including dicots such as alfalfa, almond, amaranth, apple, apricot, avocado, beans (including kidney beans, lima beans, dry beans, green beans), broccoli, cabbage, canola, carrot, cashew, castor bean, cherry, chick peas, chicory, clover, cocoa, coffee, cotton, crambe, flax, grape, grapefruit, hazelnut, hemp, jatropha, jojoba, lemon, lentils, lettuce, linseed, mango, melon (e.g., watermelon, cantaloupe), mustard, neem, olive, orange, peach, peanut, pear, peas, pepper, plum, poppy, potato, pumpkin, oilseed rape, rapeseed (high erucic acid and canola), safflower, sesame, soybean, spinach, strawberry, sugar beet, sunflower, sweet potatoes, tea, tomato, walnut, and yams, as well as monocots such as banana, barley, bluegrass, coconut, date palm, fescue, field corn, garlic, millet, oat, oil palm, onion, palm kernel oil, pineapple, popcorn, rice, rye, ryegrass, sorghum, sudangrass, sugarcane, sweet corn, switchgrass, timothy, and wheat. Brown seaweeds, green seaweeds, red seaweeds, and microalgae can also be used.

Thus, the methods and compositions described herein can be used with dicotyledonous plants belonging, for example, to the orders Apiales, Arecales, Aristochiales, Asterales, Batales, Campanulales, Capparales, Caryophyllales, Casuarinales, Celastrales, Cornales, Cucurbitales, Diapensales, Dilleniales, Dipsacales, Ebenales, Ericales, Eucomiales, Euphorbiales, Fabales, Fagales, Gentianales, Geraniales, Haloragales, Hamamelidales, Illiciales, Juglandales, Lamiales, Laurales, Lecythidales, Leitneriales, Linales, Magniolales, Malvales, Myricales, Myrtales, Nymphaeales,

Papaverales, Piperales, Plantaginales, Plumbaginales, Podostemales, Polemoniales, Polygalales, Polygonales, Primulales, Proteales, Rafflesiales, Ranunculales, Rhamniales, Rosales, Rubiales, Salicales, Santales, Sapindales, Sarraceniaceae, Scrophulariales, Solanales, Trochodendrales, Theales, Umbellales, Urticales, and Violales. The methods and compositions described herein also can be utilized with monocotyledonous plants such as those belonging to the orders Alismatales, Arales, Arecales, Asparagales, Bromeliales, Commelinales, Cyclanthales, Cyperales, Eriocaulales, Hydrocharitales, Juncales, Liliales, Najadales, Orchidales, Pandanales, Poales, Restionales, Triuridales, Typhales, Zingiberales, and with plants belonging to Gymnospermae, e.g., *Cycadales*, *Ginkgoales*, *Gnetales*, and *Pinales*.

The methods and compositions can be used over a broad range of plant species, including species from the dicot genera *Amaranthus*, *Anacardium*, *Arachis*, *Azadirachta*, *Brassica*, *Calendula*, *Camellia*, *Canarium*, *Cannabis*, *Capsicum*, *Carthainus*, *Cicer*, *Cichorium*, *Cinnamomum*, *Citrus*, *Citrus*, *Coffea*, *Corylus*, *Crambe*, *Cucumis*, *Cucurbita*, *Daucus*, *Dioscorea*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Jatropha*, *Juglans*, *Lactuca*, *Lens*, *Linum*, *Lycopersicon*, *Malus*, *Mangifera*, *Medicago*, *Mentha*, *Nicotiana*, *Ocimum*, *Olea*, *Papaver*, *Persea*, *Phaseolus*, *Pistacia*, *Pisum*, *Prunus*, *Pyrus*, *Ricinus*, *Rosmarinus*, *Salvia*, *Sesamum*, *Simmondsia*, *Solanum*, *Spinacia*, *Theobroma*, *Thymus*, *Trifolium*, *Vaccinium*, *Vigna*, and *Vitis*; and the monocot genera *Allium*, *Ananas*, *Asparagus*, *Avena*, *Cocos*, *Curcuma*, *Elaeis*, *Festuca*, *Festulolium*, *Hordeum*, *Lemna*, *Lolium*, *Musa*, *Oryza*, *Panicum*, *Pennisetum*, *Phleum*, *Poa*, *Saccharum*, *Secale*, *Sorghum*, *Triticosecale*, *Triticum*, and *Zea*; and the gymnosperm genera *Abies*, *Cunninghamia*, *Picea*, *Pinus*, *Populus*, and *Pseudotsuga*.

The methods and compositions described herein also can be used with brown seaweeds, e.g., *Ascophyllum nodosum*, *Fucus vesiculosus*, *Fucus serratus*, *Himantalia elongata*, and *Undaria pinnatifida*; red seaweeds, e.g., *Chondrus crispus*, *Cracilaria verrucosa*, *Porphyra umbilicalis*, and *Palmaria palmata*; green seaweeds, e.g., *Enteromorpha* spp. and *Ulva* spp.; and microalgae, e.g., *Spirulina* spp. (*S. platensis* and *S. maxima*) and *Odontella aurita*. In addition, the methods and compositions can be used with *Cryptocodinium cohnii*, *Schizochytrium* spp., and *Haematococcus pluvialis*.

In some embodiments, a plant is a member of the species *Arachis hypogea*, *Brassica* spp., *Carthamus tinctorius*, *Elaeis oleifera*, *Glycine max*, *Gossypium* spp., *Helianthus annuus*, *Jatropha curcas*, *Linum usitatissimum*, *Triticum aestivum*, or *Zea mays*.

#### Expression of Oil-Modulating Polypeptides

The polynucleotides and recombinant vectors described herein can be used to express an oil-modulating polypeptide in a plant species of interest. The term "expression" refers to the process of converting genetic information of a polynucleotide into RNA through transcription, which is catalyzed by an enzyme, RNA polymerase, and into protein, through translation of mRNA on ribosomes. "Up-regulation" or "activation" refers to regulation that increases the production of expression products (mRNA, polypeptide, or both) relative to basal or native states, while "down-regulation" or "repression" refers to regulation that decreases production of expression products (mRNA, polypeptide, or both) relative to basal or native states.

The polynucleotides and recombinant vectors described herein can be used to inhibit expression of an oil-modulating polypeptide in a plant species of interest. A number of nucleic acid based methods, including antisense RNA, ribozyme directed RNA cleavage, post-transcriptional gene silencing

(PTGS), e.g., RNA interference (RNAi), and transcriptional gene silencing (TGS) can be used to inhibit gene expression in plants. Antisense technology is one well-known method. In this method, a nucleic acid segment from a gene to be repressed is cloned and operably linked to a regulatory region and a transcription termination sequence so that the antisense strand of RNA is transcribed. The recombinant vector is then transformed into plants, as described herein, and the antisense strand of RNA is produced. The nucleic acid segment need not be the entire sequence of the gene to be repressed, but typically will be substantially complementary to at least a portion of the sense strand of the gene to be repressed. Generally, higher homology can be used to compensate for the use of a shorter sequence. Typically, a sequence of at least 30 nucleotides is used, e.g., at least 40, 50, 80, 100, 200, 500 nucleotides or more.

In another method, a nucleic acid can be transcribed into a ribozyme, or catalytic RNA, that affects expression of an mRNA. See, U.S. Pat. No. 6,423,885. Ribozymes can be designed to specifically pair with virtually any target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. Heterologous nucleic acids can encode ribozymes designed to cleave particular mRNA transcripts, thus preventing expression of a polypeptide. Hammerhead ribozymes are useful for destroying particular mRNAs, although various ribozymes that cleave mRNA at site-specific recognition sequences can be used. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target RNA contain a 5'-UG-3' nucleotide sequence. The construction and production of hammerhead ribozymes is known in the art. See, for example, U.S. Pat. No. 5,254,678 and WO 02/46449 and references cited therein. Hammerhead ribozyme sequences can be embedded in a stable RNA such as a transfer RNA (tRNA) to increase cleavage efficiency in vivo. Perriman et al., *Proc. Natl. Acad. Sci. USA*, 92(13): 6175-6179 (1995); de Feyter and Gaudron, *Methods in Molecular Biology*, Vol. 74, Chapter 43, "Expressing Ribozymes in Plants", Edited by Turner, P.C., Humana Press Inc., Totowa, N.J. RNA endoribonucleases which have been described, such as the one that occurs naturally in *Tetrahymena thermophila*, can be useful. See, for example, U.S. Pat. Nos. 4,987,071 and 6,423,885.

PTGS, e.g., RNAi, can also be used to inhibit the expression of a gene. For example, a construct can be prepared that includes a sequence that is transcribed into an RNA that can anneal to itself, e.g., a double stranded RNA having a stem-loop structure. In some embodiments, one strand of the stem portion of a double stranded RNA comprises a sequence that is similar or identical to the sense coding sequence of an oil-modulating polypeptide, and that is from about 10 nucleotides to about 2,500 nucleotides in length. The length of the sequence that is similar or identical to the sense coding sequence can be from 10 nucleotides to 500 nucleotides, from 15 nucleotides to 300 nucleotides, from 20 nucleotides to 100 nucleotides, or from 25 nucleotides to 100 nucleotides. The other strand of the stem portion of a double stranded RNA comprises a sequence that is similar or identical to the antisense strand of the coding sequence of the oil-modulating polypeptide, and can have a length that is shorter, the same as, or longer than the corresponding length of the sense sequence. In some cases, one strand of the stem portion of a double stranded RNA comprises a sequence that is similar or identical to the 3' or 5' untranslated region of an mRNA encoding an oil-modulating polypeptide, and the other strand of the stem portion of the double stranded RNA comprises a

sequence that is similar or identical to the sequence that is complementary to the 3' or 5' untranslated region, respectively, of the mRNA encoding the oil-modulating polypeptide. In other embodiments, one strand of the stem portion of a double stranded RNA comprises a sequence that is similar or identical to the sequence of an intron in the pre-mRNA encoding an oil-modulating polypeptide, and the other strand of the stem portion comprises a sequence that is similar or identical to the sequence that is complementary to the sequence of the intron in the pre-mRNA. The loop portion of a double stranded RNA can be from 3 nucleotides to 5,000 nucleotides, e.g., from 3 nucleotides to 25 nucleotides, from 15 nucleotides to 1,000 nucleotides, from 20 nucleotides to 500 nucleotides, or from 25 nucleotides to 200 nucleotides. The loop portion of the RNA can include an intron. A double stranded RNA can have zero, one, two, three, four, five, six, seven, eight, nine, ten, or more stem-loop structures. A construct including a sequence that is operably linked to a regulatory region and a transcription termination sequence, and that is transcribed into an RNA that can form a double stranded RNA, is transformed into plants as described herein. Methods for using RNAi to inhibit the expression of a gene are known to those of skill in the art. See, e.g., U.S. Pat. Nos. 5,034,323; 6,326,527; 6,452,067; 6,573,099; 6,753,139; and 6,777,588. See also WO 97/01952; WO 98/53083; WO 99/32619; WO 98/36083; and U.S. Patent Publications 20030175965, 20030175783, 20040214330, and 20030180945.

Constructs containing regulatory regions operably linked to nucleic acid molecules in sense orientation can also be used to inhibit the expression of a gene. The transcription product can be similar or identical to the sense coding sequence of an oil-modulating polypeptide. The transcription product can also be unpolyadenylated, lack a 5' cap structure, or contain an unsplicable intron. Methods of inhibiting gene expression using a full-length cDNA as well as a partial cDNA sequence are known in the art. See, e.g., U.S. Pat. No. 5,231,020.

In some embodiments, a construct containing a nucleic acid having at least one strand that is a template for both sense and antisense sequences that are complementary to each other is used to inhibit the expression of a gene. The sense and antisense sequences can be part of a larger nucleic acid molecule or can be part of separate nucleic acid molecules having sequences that are not complementary. The sense or antisense sequence can be a sequence that is identical or complementary to the sequence of an mRNA, the 3' or 5' untranslated region of an mRNA, or an intron in a pre-mRNA encoding an oil-modulating polypeptide. In some embodiments, the sense or antisense sequence is identical or complementary to a sequence of the regulatory region that drives transcription of the gene encoding an oil-modulating polypeptide. In each case, the sense sequence is the sequence that is complementary to the antisense sequence.

The sense and antisense sequences can be any length greater than about 12 nucleotides (e.g., 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, or more nucleotides). For example, an antisense sequence can be 21 or 22 nucleotides in length. Typically, the sense and antisense sequences range in length from about 15 nucleotides to about 30 nucleotides, e.g., from about 18 nucleotides to about 28 nucleotides, or from about 21 nucleotides to about 25 nucleotides.

In some embodiments, an antisense sequence is a sequence complementary to an mRNA sequence encoding an oil-modulating polypeptide described herein. The sense sequence complementary to the antisense sequence can be a sequence present within the mRNA of the oil-modulating

polypeptide. Typically, sense and antisense sequences are designed to correspond to a 15-30 nucleotide sequence of a target mRNA such that the level of that target mRNA is reduced.

In some embodiments, a construct containing a nucleic acid having at least one strand that is a template for more than one sense sequence (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10 or more sense sequences) can be used to inhibit the expression of a gene. Likewise, a construct containing a nucleic acid having at least one strand that is a template for more than one antisense sequence (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10 or more antisense sequences) can be used to inhibit the expression of a gene. For example, a construct can contain a nucleic acid having at least one strand that is a template for two sense sequences and two antisense sequences. The multiple sense sequences can be identical or different, and the multiple antisense sequences can be identical or different. For example, a construct can have a nucleic acid having one strand that is a template for two identical sense sequences and two identical antisense sequences that are complementary to the two identical sense sequences. Alternatively, an isolated nucleic acid can have one strand that is a template for (1) two identical sense sequences 20 nucleotides in length, (2) one antisense sequence that is complementary to the two identical sense sequences 20 nucleotides in length, (3) a sense sequence 30 nucleotides in length, and (4) three identical antisense sequences that are complementary to the sense sequence 30 nucleotides in length. The constructs provided herein can be designed to have any arrangement of sense and antisense sequences. For example, two identical sense sequences can be followed by two identical antisense sequences or can be positioned between two identical antisense sequences.

A nucleic acid having at least one strand that is a template for one or more sense and/or antisense sequences can be operably linked to a regulatory region to drive transcription of an RNA molecule containing the sense and/or antisense sequencers). In addition, such a nucleic acid can be operably linked to a transcription terminator sequence, such as the terminator of the nopaline synthase (nos) gene. In some cases, two regulatory regions can direct transcription of two transcripts: one from the top strand, and one from the bottom strand. See, for example, Yan et al., *Plant Physiol.*, 141:1508-1518 (2006). The two regulatory regions can be the same or different. The two transcripts can form double-stranded RNA molecules that induce degradation of the target RNA. In some cases, a nucleic acid can be positioned within a T-DNA or P-DNA such that the left and right T-DNA border sequences, or the left and right border-like sequences of the P-DNA, flank or are on either side of the nucleic acid. The nucleic acid sequence between the two regulatory regions can be from about 15 to about 300 nucleotides in length. In some embodiments, the nucleic acid sequence between the two regulatory regions is from about 15 to about 200 nucleotides in length, from about 15 to about 100 nucleotides in length, from about 15 to about 50 nucleotides in length, from about 18 to about 50 nucleotides in length, from about 18 to about 40 nucleotides in length, from about 18 to about 30 nucleotides in length, or from about 18 to about 25 nucleotides in length.

In some nucleic-acid based methods for inhibition of gene expression in plants, a suitable nucleic acid can be a nucleic acid analog. Nucleic acid analogs can be modified at the base moiety, sugar moiety, or phosphate backbone to improve, for example, stability, hybridization, or solubility of the nucleic acid. Modifications at the base moiety include deoxyuridine for deoxythymidine, and 5-methyl-2'-deoxycytidine and 5-bromo-2'-deoxycytidine for deoxycytidine. Modifications of the sugar moiety include modification of the 2' hydroxyl of

the ribose sugar to form 2'-O-methyl or 2'-O-allyl sugars. The deoxyribose phosphate backbone can be modified to produce morpholino nucleic acids, in which each base moiety is linked to a six-membered morpholino ring, or peptide nucleic acids, in which the deoxyphosphate backbone is replaced by a pseudopeptide backbone and the four bases are retained. See, for example, Summerton and Weller, 1997, *Antisense Nucleic Acid Drug Dev.*, 7:187-195; Hyrup et al., *Bioorgan. Med. Chem.*, 4:5-23 (1996). In addition, the deoxyphosphate backbone can be replaced with, for example, a phosphorothioate or phosphorodithioate backbone, a phosphoroamidite, or an alkyl phosphotriester backbone.

#### Transgenic Plant Phenotypes

In some embodiments, a plant in which expression of an oil-modulating polypeptide is modulated can have increased levels of seed oil. For example, an oil-modulating polypeptide described herein can be expressed in a transgenic plant, resulting in increased levels of seed oil. The seed oil level can be increased by at least 2 percent, e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, or more than 75 percent, as compared to the seed oil level in a corresponding control plant that does not express the transgene. In some embodiments, a plant in which expression of an oil-modulating polypeptide is modulated can have decreased levels of seed oil. The seed oil level can be decreased by at least 2 percent, e.g., 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, or more than 35 percent, as compared to the seed oil level in a corresponding control plant that does not express the transgene.

Plants for which modulation of levels of seed oil can be useful include, without limitation, almond, cashew, castor bean, coconut, corn, cotton, flax, hazelnut, hemp, jatropha, linseed, mustard, neem, oil palm, peanut, poppy, pumpkin, rapeseed, rice, safflower, sesame seed, soybean, sunflower, and walnut. Increases in seed oil in such plants can provide increased yields of oil extracted from the seed and increased caloric content in foodstuffs and animal feed produced from the seed. Decreases in seed oil in such plants can be useful in situations where caloric intake should be restricted.

In some embodiments, a plant in which expression of an oil-modulating polypeptide is modulated can have increased or decreased levels of oil in one or more non-seed tissues, e.g., leaf tissues, stem tissues, root or corn tissues, or fruit tissues other than seed. For example, the oil level can be increased by at least 2 percent, e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75 or more than 75 percent, as compared to the oil level in a corresponding control plant that does not express the transgene. In some embodiments, a plant in which expression of an oil-modulating polypeptide is modulated can have decreased levels of oil in one or more non-seed tissues. The oil level can be decreased by at least 2 percent, e.g., 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, or more than 35 percent, as compared to the oil level in a corresponding control plant that does not express the transgene.

Plants for which modulation of levels of oil in non-seed tissues can be useful include, without limitation, alfalfa, apple, avocado, beans, carrot, cherry, coconut, coffee, grapefruit, lemon, lettuce, oat, olive, onion, orange, palm, peach, peanut, pear, pineapple, potato, ryegrass, sudangrass, switchgrass, and tomato. Increases in non-seed oil in such plants can provide increased oil and caloric content in edible plants, including animal forage.

In some embodiments, a plant in which expression of an oil-modulating polypeptide having an amino acid sequence corresponding to SEQ ID NO:94, SEQ ID NO:81, SEQ ID NO:111, or SEQ ID NO:136 is modulated can have increased

levels of seed protein accompanying increased levels of seed oil. The protein level can be increased by at least 2 percent, e.g., 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, or 40 percent, as compared to the protein level in a corresponding control plant that does not express the transgene.

Typically, a difference (e.g., an increase) in the amount of oil or protein in a transgenic plant or cell relative to a control plant or cell is considered statistically significant at  $p \leq 0.05$  with an appropriate parametric or non-parametric statistic, e.g., Chi-square test, Student's t-test, Mann-Whitney test, or F-test. In some embodiments, a difference in the amount of oil or protein is statistically significant at  $p < 0.01$ ,  $p < 0.005$ , or  $p < 0.001$ . A statistically significant difference in, for example, the amount of oil in a transgenic plant compared to the amount in cells of a control plant indicates that the recombinant nucleic acid present in the transgenic plant results in altered oil levels.

The phenotype of a transgenic plant is evaluated relative to a control plant that does not express the exogenous polynucleotide of interest, such as a corresponding wild type plant, a corresponding plant that is not transgenic for the exogenous polynucleotide of interest but otherwise is of the same genetic background as the transgenic plant of interest, or a corresponding plant of the same genetic background in which expression of the polypeptide is suppressed, inhibited, or not induced (e.g., where expression is under the control of an inducible promoter). A plant is said "not to express" a polypeptide when the plant exhibits less than 10%, e.g., less than 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.01%, or 0.001%, of the amount of polypeptide or mRNA encoding the polypeptide exhibited by the plant of interest. Expression can be evaluated using methods including, for example, RT-PCR, Northern blots, S1 RNase protection, primer extensions, Western blots, protein gel electrophoresis, immunoprecipitation, enzyme-linked immunoassays, chip assays, and mass spectrometry. It should be noted that if a polypeptide is expressed under the control of a tissue-preferential or broadly expressing promoter, expression can be evaluated in the entire plant or in a selected tissue. Similarly, if a polypeptide is expressed at a particular time, e.g., at a particular time in development or upon induction, expression can be evaluated selectively at a desired time period.

Information that the polypeptides disclosed herein can modulate oil content can be useful in breeding of crop plants. Based on the effect of disclosed polypeptides on oil content, one can search for and identify polymorphisms linked to genetic loci for such polypeptides. Polymorphisms that can be identified include simple sequence repeats (SSRs), rapid amplification of polymorphic DNA (RAPDs), amplified fragment length polymorphisms (AFLPs) and restriction fragment length polymorphisms (RFLPs).

If a polymorphism is identified, its presence and frequency in populations analyzed to determine if it is statistically significantly correlated to an alteration in oil content. Those polymorphisms that are correlated with an alteration in oil content can be incorporated into a marker assisted breeding program to facilitate the development of lines that have a desired alteration in oil content. Typically, a polymorphism identified in such a manner is used with polymorphisms at other loci that are also correlated with a desired alteration in oil content.

#### Articles of Manufacture

Transgenic plants provided herein have particular uses in the agricultural and nutritional industries. For example, transgenic plants described herein can be used to make food products and animal feed. Suitable plants with which to make such products include almond, avocado, cashew, coconut, corn,



flax, olive, peanut, soybean, sunflower, and walnut. Such products are useful to provide desired oil and caloric content in the diet.

Transgenic plants provided herein can also be used to make vegetable oil. Vegetable oils can be chemically extracted from transgenic plants using a solvent, such as hexane. In some cases, olive, coconut and palm oils can be produced by mechanical extraction, such as expeller-pressed extraction. Oil presses, such as the screw press and the ram press, can also be used. Suitable plants from which to make oil include almond, apricot, avocado, canola, cashew, castor bean, coconut, corn, cotton, flax, grape, hazelnut, hemp, mustard, neem, olive, palm, peanut, poppy, pumpkin, rapeseed, rice, safflower, sesame, soybean, sunflower, and walnut. Such oils can be used for frying, baking, and spray coating applications. Vegetable oils also can be used to make margarine, processed foods, oleochemicals, and essential oils. Vegetable oils are used in the electrical industry as insulators. Vegetable oils are also used as lubricants. Vegetable oil derivatives can be used in the manufacture of polymers.

Vegetable oil from transgenic plants provided herein can also be used as fuel. For example, vegetable oil can be used as fuel in a vehicle that heats the oil before it enters the fuel system. Heating vegetable oil to 150° F. reduces the viscosity of the oil sufficiently for use in diesel engines, such as Mercedes-Benz® diesel engines. The viscosity of the oil can also be reduced before it enters the tank so that neither the engine or the vehicle needs modification. Methods of reducing oil viscosity include: transesterification, pyrolysis, micro emulsion, blending and thermal depolymerization. The transesterification refining process creates esters from vegetable oil by using an alcohol in the presence of a catalyst. This reaction takes a triglyceride molecule, or a complex fatty acid, neutralizes the free fatty acids and removes the glycerin, thereby creating an alcohol ester. One method of transesterification mixes methanol with sodium hydroxide and then aggressively mixes the resulting methoxide with vegetable oil, which results in a methyl ester. Ester-based oxygenated fuel made from vegetable oil is known as biodiesel. Biodiesel can be used as a pure fuel or blended with petroleum in any percentage. B5 biodiesel, for example, is a blend of 5% biodiesel and 95% petroleum diesel. B20 biodiesel, including BioWillie® diesel fuel, is produced by blending 20% biodiesel and 80% petroleum diesel.

Use of biodiesel is beneficial for the environment because it is associated with reduced emissions compared to the use of petroleum diesel. In addition, biodiesel is a biodegradable, nontoxic fuel that is made from renewable materials. Plants that can be used as sources of oil for biodiesel production include canola, cotton, flax, jatropha, oil palm, safflower, soybean, and sunflower.

Seeds of transgenic plants described herein can be conditioned and bagged in packaging material by means known in the art to form an article of manufacture. Packaging material such as paper and cloth are well known in the art. A package of seed can have a label e.g., a tag or label secured to the packaging material, a label printed on the packaging material, or a label inserted within the package. The label can indicate that plants grown from the seeds contained within the package can produce a crop having an altered level of oil relative to corresponding control plants.

The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

## EXAMPLES

## Example 1

## Transgenic Plants

The following symbols are used in the Examples: T<sub>1</sub>: first generation transformant; T<sub>2</sub>: second generation, progeny of self-pollinated T<sub>1</sub> plants; T<sub>3</sub>: third generation, progeny of self-pollinated T<sub>2</sub> plants; T<sub>4</sub>: fourth generation, progeny of self-pollinated T<sub>3</sub> plants. Independent transformations are referred to as events.

The following is a list of nucleic acids that were isolated from *Arabidopsis thaliana* plants. Ceres Clone 25429 (At3g44590; SEQ ID NO:93) is a cDNA clone that is predicted to encode a 113 amino acid (SEQ ID NO:94) ribosomal polypeptide. Ceres Clone 41573 (At5g15140; SEQ ID NO:80) is a cDNA clone that is predicted to encode a 298 amino acid (SEQ ID NO:81) aldose 1-epimerase polypeptide. Ceres Clone 5750 (At2g27960; SEQ ID NO:110) is a cDNA clone that is predicted to encode an 87 amino acid (SEQ ID NO:111) cyclin-dependent kinase regulatory subunit polypeptide. Ceres Clone 121021 (SEQ ID NO:151) is a DNA clone that is predicted to encode a 199 amino acid (SEQ ID NO:152) polypeptide. Ceres Clone 158765 (SEQ ID NO:158) is a DNA clone that is predicted to encode a 149 amino acid (SEQ ID NO:159) polypeptide. Ceres Clone 16403 (SEQ ID NO:170) is a DNA clone that is predicted to encode a 238 amino acid (SEQ ID NO:171) PsbP polypeptide. Ceres Clone 19244 (SEQ ID NO:175) is a DNA clone that is predicted to encode a 220 amino acid (SEQ ID NO:176) AP2 domain-containing polypeptide. Ceres Clone 28635 (SEQ ID NO:177) is a DNA clone that is predicted to encode a 410 amino acid (SEQ ID NO:178) squalene/phytoene synthase polypeptide. Ceres Clone 35698 (SEQ ID NO:192) is a DNA clone that is predicted to encode a 109 amino acid (SEQ ID NO:193) ubiquitin-conjugating enzyme polypeptide. Ceres Clone 36412 (SEQ ID NO:331) is a DNA clone that is predicted to encode a 358 amino acid (SEQ ID NO:332) polypeptide. Ceres Clone 368 (SEQ ID NO:341) is a DNA clone that is predicted to encode a 52 amino acid (SEQ ID NO:342) polypeptide. Ceres Clone 41046 (SEQ ID NO:343) is a DNA clone that is predicted to encode a 220 amino acid (SEQ ID NO:344) cytochrome p450 polypeptide. Ceres Clone 4829 (SEQ ID NO:345) is a DNA clone that is predicted to encode a 180 amino acid (SEQ ID NO:346) polypeptide having a DUF662 domain. Ceres Clone 5426 (SEQ ID NO:358) is a DNA clone that is predicted to encode a 143 amino acid (SEQ ID NO:359) ribosomal L28e protein family polypeptide. Ceres Clone 7894 (SEQ ID NO:373) is a DNA clone that is predicted to encode a 512 amino acid (SEQ ID NO:374) Major Facilitator Superfamily transporter polypeptide. Ceres Clone 8161 (SEQ ID NO:397) is a cDNA clone that is predicted to encode a 218 amino acid (SEQ ID NO:398) polypeptide.

Ceres Clone 218626 (SEQ ID NO:135) is a cDNA clone that was isolated from *Zea mays* and that is predicted to encode a 454 amino acid (SEQ ID NO:136) tryptophan/tyrosine permease family polypeptide.

Each isolated nucleic acid described above was cloned into a Ti plasmid vector, CRS 338, containing a phosphinothricin acetyltransferase gene which confers Finale™ resistance to transformed plants. Constructs were made using CRS 338 that contained Ceres Clone 25429, Ceres Clone 41573, Ceres Clone 5750, Ceres Clone 218626, Ceres Clone 121021, Ceres Clone 158765, Ceres Clone 16403, Ceres Clone 19244, Ceres Clone 28635, Ceres Clone 35698, Ceres Clone 36412, Ceres

Clone 368, Ceres Clone 41046, Ceres Clone 4829, Ceres Clone 5426, Ceres Clone 7894, or Ceres Clone 8161, each operably linked to a CaMV 35S promoter. Wild-type *Arabidopsis thaliana* ecotype Wassilewskija (Ws) plants were transformed separately with each construct. The transformations were performed essentially as described in Bechtold et al., *C. R. Acad. Sci. Paris*, 316:1194-1199 (1993).

Transgenic *Arabidopsis* lines containing Ceres Clone 25429, Ceres Clone 41573, Ceres Clone 5750, Ceres Clone 218626, Ceres Clone 5426, Ceres Clone 35698, Ceres Clone 4829, Ceres Clone 28635, Ceres Clone 16403, Ceres Clone 41046, Ceres Clone 19244, Ceres Clone 7894, Ceres Clone 158765, Ceres Clone 121021, Ceres Clone 36412, Ceres Clone 8161, or Ceres Clone 368 were designated ME01597, ME01720, ME01833, ME02065, ME01902, ME00072, ME00085, ME00147, ME00896, ME00900, ME00913, ME01704, ME02505, ME02525, ME00902, ME00914, or ME01754, respectively. The presence of each vector containing a Ceres clone described above in the respective transgenic *Arabidopsis* line transformed with the vector was confirmed by Finale™ resistance, polymerase chain reaction (PCR) amplification from green leaf tissue extract, and/or sequencing of PCR products. As controls, wild-type *Arabidopsis* ecotype Ws plants were transformed with the empty vector CRS 338.

The physical appearances of T<sub>1</sub> plants from ten events each of ME01597 and ME01720 were similar to those of corresponding control plants. The physical appearances of T<sub>1</sub> plants from nine out of ten events of ME01833 were comparable to those of the controls. Event-08 of ME01833 was taller and had reduced fertility compared to control plants. Event-08 of ME02065 also had reduced fertility compared to control plants. The physical appearances of T<sub>1</sub> plants from nine additional events of ME02065 were similar to those of corresponding control plants.

### Example 2

#### Analysis of Oil Content in Transgenic *Arabidopsis* Seeds

An analytical method based on Fourier transform near-infrared (FT-NIR) spectroscopy was developed, validated, and used to perform a high-throughput screen of transgenic seed lines for alterations in seed oil content. To calibrate the FT-NIR spectroscopy method, a sub-population of transgenic seed lines was randomly selected and analyzed for oil content using a direct primary method. Fatty acid methyl ester (FAME) analysis by gas chromatography-mass spectroscopy (GC-MS) was used as the direct primary method to determine the total fatty acid content for each seed line and produce the FT-NIR spectroscopy calibration curves for oil.

To analyze seed oil content using GC-MS, seed tissue was homogenized in liquid nitrogen using a mortar and pestle to create a powder. The tissue was weighed, and 5.0±0.25 mg were transferred into a 2 mL Eppendorf tube. The exact weight of each sample was recorded. One mL of 2.5% H<sub>2</sub>SO<sub>4</sub> (v/v in methanol) and 20 µL of undecanoic acid internal standard (1 mg/mL in hexane) were added to the weighed seed tissue. The tubes were incubated for two hours at 90° C. in a pre-equilibrated heating block. The samples were removed from the heating block and allowed to cool to room temperature. The contents of each Eppendorf tube were poured into a 15 mL polypropylene conical tube, and 1.5 mL of a 0.9% NaCl solution and 0.75 mL of hexane were added to each tube. The tubes were vortexed for 30 seconds and incubated at room temperature for 15 minutes. The samples

were then centrifuged at 4,000 rpm for 5 minutes using a bench top centrifuge. If emulsions remained, then the centrifugation step was repeated until they were dissipated. One hundred µL of the hexane (top) layer was pipetted into a 1.5 mL autosampler vial with minimum volume insert. The samples were stored no longer than 1 week at -80° C. until they were analyzed.

Samples were analyzed using a Shimadzu QP-2010 GC-MS (Shimadzu Scientific Instruments, Columbia, Md.). The first and last sample of each batch consisted of a blank (hexane). Every fifth sample in the batch also consisted of a blank. Prior to sample analysis, a 7-point calibration curve was generated using the Supelco 37 component FAME mix (0.00004 mg/mL to 0.2 mg/mL). The injection volume was 1 µL.

The GC parameters were as follows: column oven temperature: 70° C., inject temperature: 230° C., inject mode: split, flow control mode: linear velocity, column flow: 1.0 mL/min, pressure: 53.5 mL/min, total flow: 29.0 mL/min, purge flow: 3.0 mL/min, split ratio: 25.0. The temperature gradient was as follows: 70° C. for 5 minutes, increasing to 350° C. at a rate of 5 degrees per minute, and then held at 350° C. for 1 minute. The MS parameters were as follows: ion source temperature: 200° C., interface temperature: 240° C., solvent cut time: 2 minutes, detector gain mode: relative, detector gain: 0.6 kV, threshold: 1000, group: 1, start time: 3 minutes, end time: 62 minutes, ACQ mode: scan, interval: 0.5 second, scan speed: 666 amu/sec., start M/z: 40, end M/z: 350. The instrument was tuned each time the column was cut or a new column was used.

The data were analyzed using the Shimadzu GC-MS Solutions software. Peak areas were integrated and exported to an Excel spreadsheet. Fatty acid peak areas were normalized to the internal standard, the amount of tissue weighed, and the slope of the corresponding calibration curve generated using the FAME mixture. Peak areas were also multiplied by the volume of hexane (0.75 mL) used to extract the fatty acids.

The same seed lines that were analyzed using GC-MS were also analyzed by FT-NIR spectroscopy, and the oil values determined by the GC-MS primary method were entered into the FT-NIR chemometrics software (Bruker Optics, Billerica, Mass.) to create a calibration curve for oil content. The actual oil content of each seed line analyzed using GC-MS was plotted on the x-axis of the calibration curve. The y-axis of the calibration curve represented the predicted values based on the best-fit line. Data points were continually added to the calibration curve data set.

T<sub>2</sub> seed from each transgenic plant line was analyzed by FT-NIR spectroscopy. Sarstedt tubes containing seeds were placed directly on the lamp, and spectra were acquired through the bottom of the tube. The spectra were analyzed to determine seed oil content using the FT-NIR chemometrics software (Bruker Optics) and the oil calibration curve. Results for experimental samples were compared to population means and standard deviations calculated for transgenic seed lines that were planted within 30 days of the lines being analyzed and grown under the same conditions. Typically, results from three to four events of each of 400 to 1600 different transgenic lines were used to calculate a population mean. Each data point was assigned a z-score ( $z=(x-\text{mean})/\text{std}$ ), and a p-value was calculated for the z-score.

Transgenic seed lines with oil levels in T<sub>2</sub> seed that differed by more than two standard deviations from the population mean were selected for evaluation of oil levels in the T<sub>3</sub> generation. All events of selected lines were planted in individual pots. The pots were arranged randomly in flats along with pots containing matched control plants in order to mini-

mize microenvironment effects. Matched control plants contained an empty version of the vector used to generate the transgenic seed lines. T<sub>3</sub> seed from up to five plants from each event was collected and analyzed individually using FT-NIR spectroscopy. Data from replicate samples were averaged and compared to controls using the Student's t-test.

### Example 3

#### Analysis of Protein Content in Transgenic *Arabidopsis* Seeds

An analytical method based on Fourier transform near-infrared (FT-NIR) spectroscopy was developed, validated, and used to perform a high-throughput screen of transgenic seed lines for alterations in seed protein content. To calibrate the FT-NIR spectroscopy method, total nitrogen elemental analysis was used as a primary method to analyze a sub-population of randomly selected transgenic seed lines. The overall percentage of nitrogen in each sample was determined. Percent nitrogen values were multiplied by a conversion factor to obtain percent total protein values. A conversion factor of 5.30 was selected based on data for cotton, sunflower, safflower, and sesame seed (Rhee, K. C., Determination of Total Nitrogen In Handbook of Food Analytical Chemistry—Water, Proteins, Enzymes, Lipids, and Carbohydrates (R. Wrolstad et al., ed.), John Wiley and Sons, Inc., p. 105, (2005)). The same seed lines were then analyzed by FT-NIR spectroscopy, and the protein values calculated via the primary method were entered into the FT-NIR chemometrics software (Bruker Optics, Billerica, Mass.) to create a calibration curve for analysis of seed protein content by FT-NIR spectroscopy.

Elemental analysis was performed using a FlashEA 1112 NC Analyzer (Thermo Finnigan, San Jose, Calif.). To analyze total nitrogen content, 2.00±0.15 mg of dried transgenic *Arabidopsis* seed was weighed into a tared tin cup. The tin cup with the seed was weighed, crushed, folded in half, and placed into an autosampler slot on the FlashEA 1112 NC Analyzer (Thermo Finnigan). Matched controls were prepared in a manner identical to the experimental samples and spaced evenly throughout the batch. The first three samples in every batch were a blank (empty tin cup), a bypass, (approximately 5 mg of aspartic acid), and a standard (5.00±0.15 mg aspartic acid), respectively. Blanks were entered between every 15 experimental samples. Each sample was analyzed in triplicate.

The FlashEA 1112 NC Analyzer (Thermo Finnigan) instrument parameters were as follows: left furnace 900° C., right furnace 840° C., oven 50° C., gas flow carrier 130 mL/min., and gas flow reference 100 mL/min. The data parameter LLOD was 0.25 mg for the standard and different for other materials. The data parameter LLOQ was 3.0 mg for the standard, 1.0 mg for seed tissue, and different for other materials.

Quantification was performed using the Eager 300 software (Thermo Finnigan). Replicate percent nitrogen measurements were averaged and multiplied by a conversion factor of 5.30 to obtain percent total protein values. For results to be considered valid, the standard deviation between replicate samples was required to be less than 10%. The percent nitrogen of the aspartic acid standard was required to be within ±1.0% of the theoretical value. For a run to be declared valid, the weight of the aspartic acid (standard) was required to be between 4.85 and 5.15 mg, and the blank(s) were required to have no recorded nitrogen content.

The same seed lines that were analyzed for elemental nitrogen content were also analyzed by FT-NR spectroscopy, and the percent total protein values determined by elemental analysis were entered into the FT-NIR chemometrics software (Bruker Optics, Billerica, Mass.) to create a calibration curve for protein content. The protein content of each seed line based on total nitrogen elemental analysis was plotted on the x-axis of the calibration curve. The y-axis of the calibration curve represented the predicted values based on the best-fit line. Data points were continually added to the calibration curve data set.

T<sub>2</sub> seed from each transgenic plant line was analyzed by FT-NIR spectroscopy. Sarstedt tubes containing seeds were placed directly on the lamp, and spectra were acquired through the bottom of the tube. The spectra were analyzed to determine seed protein content using the FT-NIR chemometrics software (Bruker Optics) and the protein calibration curve. Results for experimental samples were compared to population means and standard deviations calculated for transgenic seed lines that were planted within 30 days of the lines being analyzed and grown under the same conditions. Typically, results from three to four events of each of 400 to 1600 different transgenic lines were used to calculate a population mean. Each data point was assigned a z-score ( $z=(x-\text{mean})/\text{std}$ ), and a p-value was calculated for the z-score.

Transgenic seed lines with oil levels in T<sub>2</sub> seed that differed by more than two standard deviations from the population mean were also analyzed to determine protein levels in the T<sub>3</sub> generation. Events of selected lines were planted in individual pots. The pots were arranged randomly in flats along with pots containing matched control plants in order to minimize microenvironment effects. Matched control plants contained an empty version of the vector used to generate the transgenic seed lines. T<sub>3</sub> seed from up to five plants from each event was collected and analyzed individually using FT-NIR spectroscopy. Data from replicate samples were averaged and compared to controls using the Student's t-test.

### Example 4

#### Results for ME01597 Events

T<sub>2</sub> and T<sub>3</sub> seed from three events and four events, respectively, of ME01597 containing Ceres Clone 25429 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from two events of ME01597 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME01597. As presented in Table 1, the oil content was increased to 123% and 124% in seed from events-03 and -06, respectively, compared to the population mean.

TABLE 1

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME01597 events containing Ceres Clone 25429					
	Event - 01	Event - 03	Event - 06	Event - 09	Control
Oil content (% control) in T <sub>2</sub> seed	101	123	124	No data	100 ± 11*
p-value**	0.32	0.04	0.03	No data	N/A
Oil content (% control) in T <sub>3</sub> seed	106 ± 2	107 ± 4	107 ± 3	99 ± 4	100 ± 4

59

TABLE 1-continued

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME01597 events containing Ceres Clone 25429					
	Event - 01	Event - 03	Event - 06	Event - 09	Control
p-value***	<0.01	0.05	<0.01	0.70	N/A
No. of T <sub>2</sub> plants	5	4	4	5	31

\*Population mean of the oil content of seed from transgenic lines planted within 30 days of ME01597. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from three events of ME01597 was significantly increased compared to the oil content of corresponding control seeds. As presented in Table 1, the oil content was increased to 106% in seed from event-01 and to 107% in seed from events-03 and -06 compared to the oil content in control seed.

T<sub>2</sub> and T<sub>3</sub> seed from three events and four events, respectively, of ME01597 containing Ceres Clone 25429 was also analyzed for protein content using FT-NIR spectroscopy as described in Example 3.

The protein content in T<sub>2</sub> seed from ME01597 events was not observed to differ significantly from the mean protein content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME01597 (Table 2).

TABLE 2

Protein content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME01597 events containing Ceres Clone 25429					
	Event - 01	Event - 03	Event - 06	Event - 09	Control
Protein content (% control) in T <sub>2</sub> seed	98	98	93	No data	100 ± 9*
p-value**	0.40	0.29	0.27	No data	N/A
Protein content (% control) in T <sub>3</sub> seed	109 ± 2	110 ± 4	106 ± 7	110 ± 3	100 ± 5
p-value***	<0.01	0.01	0.15	<0.01	N/A
No. of T <sub>2</sub> plants	5	4	4	5	31

\*Population mean of the protein content in seed from transgenic lines planted within 30 days of ME01597. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The protein content in T<sub>3</sub> seed from three events of ME01597 was significantly increased compared to the protein content in corresponding control seed. As presented in Table 2, the protein content was increased to 109% in seed from event-01 and to 110% in seed from events-03 and -09 compared to the protein content in control seed.

There were no observable or statistically significant differences between T<sub>2</sub> ME01597 and control plants in germination, onset of flowering, rosette area, fertility, and general morphology/architecture.

### Example 5

#### Results for ME01720 Events

T<sub>2</sub> and T<sub>3</sub> seed from five events of ME01720 containing Ceres Clone 41573 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from three events of ME01720 was significantly increased compared to the mean oil content

60

in seed from transgenic *Arabidopsis* lines planted within 30 days of ME01720. As presented in Table 3, the oil content was increased to 123% in seed from events-01 and -03 and to 130% in seed from event-06 compared to the population mean.

TABLE 3

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME01720 events containing Ceres Clone 41573						
	Event - 01	Event - 02	Event - 03	Event - 04	Event - 06	Control
Oil content (% control) in T <sub>2</sub> seed	123	98	123	112	130	100 ± 11*
p-value**	0.04	0.31	0.04	0.18	0.01	N/A
Oil content (% control) in T <sub>3</sub> seed	105 ± 3	100 ± 3	108 ± 1	99 ± 1	105 ± 2	100 ± 4
p-value***	0.01	1.00	<0.01	0.60	0.04	N/A
No. of T <sub>2</sub> plants	5	5	4	3	3	31

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME01720. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from three events of ME01720 was significantly increased compared to the oil content in corresponding control seeds. As presented in Table 3, the oil content was increased to 105% in seed from events-01 and -06 and to 108% in seed from event-03 compared to the oil content in control seed.

T<sub>2</sub> and T<sub>3</sub> seed from five events of ME01720 containing Ceres Clone 41573 was also analyzed for total protein content using FT-NIR spectroscopy as described in Example 3.

The protein content in T<sub>2</sub> seed from ME01720 events was not observed to differ significantly from the mean protein content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME01720 (Table 4).

TABLE 4

Protein content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME01720 events containing Ceres Clone 41573						
	Event - 01	Event - 02	Event - 03	Event - 04	Event - 06	Control
Protein content (% control) in T <sub>2</sub> seed	92	103	89	96	84	100 ± 8*
p-value**	0.24	0.40	0.15	0.36	0.06	N/A
Protein content (% control) in T <sub>3</sub> seed	107 ± 2	114 ± 3	106 ± 5	107 ± 0	106 ± 2	100 ± 5
p-value***	<0.01	<0.01	0.07	<0.01	0.01	N/A
No. of T <sub>2</sub> plants	5	5	4	3	3	31

\*Population mean of the protein content in seed from transgenic lines planted within 30 days of ME01720. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The protein content in T<sub>3</sub> seed from four events of ME01720 was significantly increased compared to the protein content in corresponding control seed. As presented in Table 4, the protein content was increased to 107%, 114%, 107%, and 106% in events-01, -02, -04, and -06, respectively, compared to the protein content in control seed.

There were no observable or statistically significant differences between T<sub>2</sub> ME01720 and control plants in germina-

## 61

tion, onset of flowering, rosette area, and fertility. The general morphology/architecture of the plants appeared wild-type in all instances except for event-01, which had a small (<30%), but statistically significant ( $p < 0.05$ ), change in plant size.

## Example 6

## Results for ME01833 Events

$T_2$  and  $T_3$  seed from four events of ME01833 containing Ceres Clone 5750 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in  $T_2$  seed from four events of ME01833 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME01833. As presented in Table 5, the oil content was increased to 135%, 138%, 143%, and 172% in seed from events-01, -02, -03, and -04, respectively, compared to the population mean.

TABLE 5

Oil content (% control) in $T_2$ and $T_3$ seed from ME01833 events containing Ceres Clone 5750					
	Event - 01	Event - 02	Event - 03	Event - 04	Control
Oil content (% control) in $T_2$ seed	135	138	143	172	100 ± 11*
p-value**	<0.01	<0.01	<0.01	<0.01	N/A
Oil content (% control) in $T_3$ seed	95 ± 4	102 ± 4	105 ± 1	112 ± 1	100 ± 4
p-value***	0.08	0.35	<0.01	<0.01	N/A
No. of $T_2$ plants	4	4	3	4	31

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME01833. Variation is presented as the standard error of the mean.

\*\*The p-values for  $T_2$  seed were calculated using z-scores.

\*\*\*The p-values for  $T_3$  seed were calculated using a Student's t-test.

The oil content in  $T_3$  seed from two events of ME01833 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 5, the oil content was increased to 105% and 112% in seed from events-03 and -04, respectively, compared to the oil content in control seed.

$T_2$  and  $T_3$  seed from four events of ME01833 containing Ceres Clone 5750 was also analyzed for total protein content using FT-NIR spectroscopy as described in Example 3.

The protein content in  $T_2$  seed from one event of ME01833 was significantly increased compared to the mean protein content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME01833. As presented in Table 6, the protein content was increased to 123% in event-04 compared to the population mean.

TABLE 6

Protein content (% control) in $T_2$ and $T_3$ seed from ME01833 events containing Ceres Clone 5750					
	Event - 01	Event - 02	Event - 03	Event - 04	Control
Protein content (% control) in $T_2$ seed	94	101	98	123	100 ± 8*
p-value**	0.32	0.41	0.40	0.01	N/A
Protein content (% control) in $T_3$ seed	109 ± 8	102 ± 3	99 ± 1	112 ± 5	100 ± 4

## 62

TABLE 6-continued

Protein content (% control) in $T_2$ and $T_3$ seed from ME01833 events containing Ceres Clone 5750					
	Event - 01	Event - 02	Event - 03	Event - 04	Control
p-value***	0.11	0.28	0.28	0.02	N/A
No. of $T_2$ plants	4	4	3	4	31

\*Population mean of the protein content in seed from transgenic lines planted within 30 days of ME01833. Variation is presented as the standard error of the mean.

\*\*The p-values for  $T_2$  seed were calculated using z-scores.

\*\*\*The p-values for  $T_3$  seed were calculated using a Student's t-test.

The protein content in  $T_3$  seed from one event of ME01833 was significantly increased compared to the protein content in corresponding control seed. As presented in Table 6, the protein content was increased to 112% in seed from event-04 compared to the protein content in control seed.

There were no observable or statistically significant differences between  $T_2$  ME01833 and control plants in germination, onset of flowering, rosette area, fertility, and general morphology/architecture.

## Example 7

## Results for ME02065 Events

$T_2$  and  $T_3$  seed from four events of ME02065 containing Ceres Clone 218626 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in  $T_2$  seed from three events of ME02065 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME02065. As presented in Table 7, the oil content was increased to 123%, 132%, and 145% in seed from events-01, -04, and -05, respectively, compared to the population mean.

TABLE 7

Oil content (% control) in $T_2$ and $T_3$ seed from ME02065 events containing Ceres Clone 218626					
	Event - 01	Event - 02	Event - 04	Event - 05	Control
Oil content (% control) in $T_2$ seed	123	115	132	145	100 ± 11*
p-value**	0.04	0.13	<0.01	<0.01	N/A
Oil content (% control) in $T_3$ seed	109 ± 4	103 ± 1	112 ± 1	102 ± 5	100 ± 4
p-value***	0.01	0.05	<0.01	0.46	N/A
No. of $T_2$ plants	4	2	5	4	31

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME02065. Variation is presented as the standard error of the mean.

\*\*The p-values for  $T_2$  seed were calculated using z-scores.

\*\*\*The p-values for  $T_3$  seed were calculated using a Student's t-test.

The oil content in  $T_3$  seed from three events of ME02065 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 7, the oil content was increased to 109%, 103%, and 112% in seed from events-01, -02, and -04, respectively, compared to the oil content in control seed.

$T_2$  and  $T_3$  seed from four events of ME02065 containing Ceres Clone 218626 was also analyzed for total protein content using FT-NIR spectroscopy as described in Example 3.

The protein content in  $T_2$  seed from ME02065 events was not observed to differ significantly from the mean protein

content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME02065 (Table 8).

TABLE 8

Protein content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME02065 events containing Ceres Clone 218626					
	Event - 01	Event - 02	Event - 04	Event - 05	Control
Protein content (% control) in T <sub>2</sub> seed	103	104	104	105	100 ± 8*
p-value**	0.39	0.37	0.37	0.34	N/A
Protein content (% control) in T <sub>3</sub> seed	105 ± 3	110 ± 1	109 ± 2	113 ± 4	100 ± 5
p-value***	0.01	<0.01	<0.01	<0.01	N/A
No. of T <sub>2</sub> plants	4	2	5	4	31

\*Population mean of the protein content in seed from transgenic lines planted within 30 days of ME02065. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The protein content in T<sub>3</sub> seed from four events of ME02065 was significantly increased compared to the protein content in corresponding control seed. As presented in Table 8, the protein content was increased to 105%, 110%, 109%, and 113% in seed from events-01, -02, -04, and -05, respectively, compared to the protein content in control seed.

There were no observable or statistically significant differences between T<sub>2</sub> ME02065 and control plants in germination, onset of flowering, rosette area, fertility, and general morphology/architecture.

### Example 8

#### Results for ME01902 Events

T<sub>2</sub> and T<sub>3</sub> seed from five events and four events, respectively, of ME01902 containing Ceres Clone 5426 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from three events of ME01902 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME01902. As presented in Table 9, the oil content was increased to 118%, 125%, and 124% in seed from events-02, -04, and -07, respectively, compared to the population mean.

TABLE 9

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME01902 events containing Ceres Clone 5426						
	Event - 01	Event - 02	Event - 03	Event - 04	Event - 07	Control
Oil content (% control) in T <sub>2</sub> seed	102	118	107	125	124	100*
p-value**	0.17	0.03	0.14	0.01	0.01	N/A
Oil content (% control) in T <sub>3</sub> seed	No data	108 ± 4	103 ± 1	104 ± 4	109 ± 5	100
p-value***	No data	0.02	<0.01	0.10	0.03	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME01902. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from three events of ME01902 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 9, the oil content was increased to 108%, 103%, and 109% in seed from events-02, -03, and -07, respectively, compared to the oil content in control seed.

T<sub>2</sub> and T<sub>3</sub> seed from five events and four events, respectively, of ME01902 containing Ceres Clone 5426 was also analyzed for total protein content using FT-NIR spectroscopy as described in Example 3.

The protein content in T<sub>2</sub> seed from ME01902 events was not observed to differ significantly from the mean protein content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME01902 (Table 10).

TABLE 10

Protein content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME01902 events containing Ceres Clone 5426						
	Event - 01	Event - 02	Event - 03	Event - 04	Event - 07	Control
Protein content (% control) in T <sub>2</sub> seed	99	106	90	93	96	100*
p-value**	0.26	0.18	0.08	0.15	0.22	N/A
Protein content (% control) in T <sub>3</sub> seed	No data	116 ± 4	113	118 ± 2	111 ± 1	100 ± 1
p-value***	No data	0.02	<0.01	<0.01	<0.01	N/A

\*Population mean of the protein content in seed from transgenic lines planted within 30 days of ME01902. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The protein content in T<sub>3</sub> seed from four events of ME01902 was significantly increased compared to the protein content in corresponding control seed. As presented in Table 10, the protein content was increased to 116%, 113%, 118%, and 111% in seed from events-02, -03, -04, and -07, respectively, compared to the protein content in control seed.

### Example 9

#### Results for ME00072 Events

T<sub>2</sub> and T<sub>3</sub> seed from seven events and five events, respectively, of ME00072 containing Ceres Clone 35698 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from three events of ME00072 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME00072. As presented in Table 11, the oil content was increased to 115%, 121%, and 117% in seed from events-02, -03, and -06, respectively, compared to the population mean.

TABLE 11

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME00072 events containing Ceres Clone 35698								
	Event - 02	Event - 03	Event - 05	Event - 06	Event - 07	Event - 08	Event - 09	Control
Oil content (% control) in T <sub>2</sub> seed	115	121	109	117	104	107	102	100 ± 12*
p-value**	0.04	0.01	0.11	0.02	0.17	0.14	0.19	N/A
Oil content (% control) in T <sub>3</sub> seed	100 ± 1	102 ± 2	101 ± 4	105 ± 1	101 ± 2	No data	No data	100 ± 4
p-value***	0.72	0.18	0.79	0.01	0.60	No data	No data	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME00072. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from one event of ME00072 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 11, the oil content was increased to 105% in seed from event-06 compared to the oil content in control seed.

#### Example 10

##### Results for ME00085 Events

T<sub>2</sub> and T<sub>3</sub> seed from six events and four events, respectively, of ME00085 containing Ceres Clone 4829 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from two events of ME00085 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME00085. As presented in Table 12, the oil content was increased to 120% and 117% in seed from events-01 and -02, respectively, compared to the population mean.

TABLE 12

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME00085 events containing Ceres Clone 4829							
	Event - 01	Event - 02	Event - 03	Event - 04	Event - 05	Event - 06	Control
Oil content (% control) in T <sub>2</sub> seed	120	117	92	98	100	109	100 ± 12*
p-value**	0.01	0.03	0.13	0.19	0.19	0.11	N/A
Oil content (% control) in T <sub>3</sub> seed	95 ± 4	111 ± 2	111 ± 2	104 ± 4	No data	No data	100 ± 4
p-value***	0.20	<0.01	<0.01	0.09	No data	No data	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME00085. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from two events of ME00085 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 12, the oil content was increased to 111% in seed from events-02 and -03 compared to the oil content in control seed.

#### Example 11

##### Results for ME00147 Events

T<sub>2</sub> and T<sub>3</sub> seed from seven events and five events, respectively, of ME00147 containing Ceres Clone 28635 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from three events of ME00147 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME00147. As presented in Table 13, the oil content was increased to 119%, 120%, and 118% in seed from events-02, -03, and -04, respectively, compared to the population mean.

TABLE 13

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME00147 events containing Ceres Clone 28635									
	Event -02	Event -03	Event -04	Event -05	Event -06	Event -07	Event -08	Event -09	Control
Oil content (% control) in T <sub>2</sub> seed	119	120	118	No data	95	108	105	102	100 ± 12*
p-value**	0.02	0.01	0.02	No data	0.16	0.13	0.16	0.19	N/A
Oil content (% control) in T <sub>3</sub> seed	102 ± 2	101 ± 0	105 ± 2	99 ± 2	97 ± 4	No data	No data	No data	100 ± 4
p-value***	0.15	0.47	0.04	0.52	0.21	No data	No data	No data	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME00147. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from one event of ME00147 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 13, the oil content was increased to 105% in seed from event-04 compared to the oil content in control seed.

#### Example 12

##### Results for ME00896 Events

T<sub>2</sub> and T<sub>3</sub> seed from three events and four events, respectively, of ME00896 containing Ceres Clone 16403 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from three events of ME00896 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME00896. As presented in Table 14, the oil content was increased to 130%, 126%, and 118% in seed from events-03, -04, and -05, respectively, compared to the population mean.

TABLE 14

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME00896 events containing Ceres Clone 16403					
	Event -02	Event -03	Event -04	Event -05	Control
Oil content (% control) in T <sub>2</sub> seed	No data	130	126	118	100 ± 10*
p-value**	No data	0.01	0.02	0.05	N/A
Oil content (% control) in T <sub>3</sub> seed	102 ± 2	100 ± 1	104 ± 2	100 ± 1	100 ± 4
p-value***	0.09	0.91	0.03	0.93	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME00896. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from one event of ME00896 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 14, the oil content was increased to 104% in seed from event-04 compared to the oil content in control seed.

#### Example 13

##### Results for ME00900 Events

T<sub>2</sub> and T<sub>3</sub> seed from three events and two events, respectively, of ME00900 containing Ceres Clone 41046 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from two events of ME00900 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME00900. As presented in Table 15, the oil content was increased to 132% and 128% in seed from events-01 and -02, respectively, compared to the population mean.

TABLE 15

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME00900 events containing Ceres Clone 41046				
	Event -01	Event -02	Event -03	Control
Oil content (% control) in T <sub>2</sub> seed	132	128	108	100 ± 10*
p-value**	0.01	0.01	0.11	N/A
Oil content (% control) in T <sub>3</sub> seed	105 ± 3	No data	97 ± 2	100 ± 4
p-value***	0.01	No data	0.06	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME00900. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from one event of ME00900 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 15, the oil content was increased to 105% in seed from event-01 compared to the oil content in control seed.

#### Example 14

##### Results for ME00913 Events

T<sub>2</sub> and T<sub>3</sub> seed from four events of ME00913 containing Ceres Clone 19244 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from two events of ME00913 was significantly increased compared to the mean oil content



69

in seed from transgenic *Arabidopsis* lines planted within 30 days of ME00913. As presented in Table 16, the oil content was increased to 130% and 128% in seed from events-01 and -04, respectively, compared to the population mean.

TABLE 16

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME00913 events containing Ceres Clone 19244					
	Event -01	Event -02	Event -03	Event -04	Control
Oil content (% control) in T <sub>2</sub> seed	130	114	91	128	100 ± 10*
p-value**	0.01	0.08	0.11	0.01	N/A
Oil content (% control) in T <sub>3</sub> seed	99 ± 5	107 ± 0	104 ± 2	106 ± 4	100 ± 4
p-value***	0.91	<0.01	0.01	0.02	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME00913. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from three events of ME00913 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 16, the oil content was increased to 107%, 104%, and 106% in seed from events-02, -03, and -04, respectively, compared to the oil content in control seed.

## Example 15

## Results for ME01704 Events

T<sub>2</sub> and T<sub>3</sub> seed from four events and five events, respectively, of ME01704 containing Ceres Clone 7894 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from three events of ME01704 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME01704. As presented in Table 17, the oil content

70

was increased to 126%, 118%, and 132% in seed from events-02, -04, and -06, respectively, compared to the population mean.

TABLE 17

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME01704 events containing Ceres Clone 7894						
	Event -01	Event -02	Event -03	Event -04	Event -06	Control
Oil content (% control) in T <sub>2</sub> seed	103	126	No data	118	132	100 ± 9*
p-value**	0.17	0.01	No data	0.04	<0.01	N/A
Oil content (% control) in T <sub>3</sub> seed	101 ± 1	103 ± 2	97 ± 6	109 ± 3	107 ± 3	100 ± 4
p-value***	0.60	0.12	0.49	<0.01	<0.01	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME01704. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from two events of ME01704 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 17, the oil content was increased to 109% and 107% in seed from events-04 and -06, respectively, compared to the oil content in control seed.

## Example 16

## Results for ME02505 Events

T<sub>2</sub> and T<sub>3</sub> seed from five events of ME02505 containing Ceres Clone 158765 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from three events of ME02505 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME02505. As presented in Table 18, the oil content was increased to 120%, 129%, and 140% in seed from events-01, -02, and -03, respectively, compared to the population mean.

TABLE 18

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME02505 events containing Ceres Clone 158765						
	Event -01	Event -02	Event -03	Event -04	Event -05	Control
Oil content (% control) in T <sub>2</sub> seed	120	129	140	104	105	100 ± 11*
p-value**	0.02	<0.01	<0.01	0.19	0.19	N/A
Oil content (% control) in T <sub>3</sub> seed	101 ± 1	100 ± 1	103 ± 1	108 ± 1	105 ± 2	100 ± 4
p-value***	0.47	0.79	0.01	<0.01	0.02	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME02505. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

## 71

The oil content in T<sub>3</sub> seed from three events of ME02505 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 18, the oil content was increased to 103%, 108%, and 105% in seed from events-03, -04, and -05, respectively, compared to the oil content in control seed.

## Example 17

## Results for ME02525 Events

T<sub>2</sub> and T<sub>3</sub> seed from four events of ME02525 containing Ceres Clone 121021 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from two events of ME02525 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME02525. As presented in Table 19, the oil content was increased to 134% and 145% in seed from events-01 and -02, respectively, compared to the population mean.

TABLE 19

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME02525 events containing Ceres Clone 121021						
	Event -01	Event -02	Event -03	Event -04	Event -07	Control
Oil content (% control) in T <sub>2</sub> seed	134	145	113	No data	103	100 ± 11*
p-value**	<0.01	<0.01	0.09	No data	0.20	N/A
Oil content (% control) in T <sub>3</sub> seed	102 ± 2	99 ± 1	98 ± 1	103 ± 2	No data	100 ± 4
p-value***	0.19	0.16	0.10	0.05	No data	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME02525. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from one event of ME02525 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 19, the oil content was increased to 103% in seed from event-04 compared to the oil content in control seed.

## Example 18

## Results for ME00902 Events

T<sub>2</sub> and T<sub>3</sub> seed from four events of ME00902 containing Ceres Clone 36412 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from ME00902 events was not observed to differ significantly from the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME00902 (Table 20).

TABLE 20

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME00902 events containing Ceres Clone 36412					
	Event -01	Event -03	Event -04	Event -05	Control
Oil content (% control) in T <sub>2</sub> seed	115	114	114	113	100 ± 10*
p-value**	0.07	0.07	0.08	0.08	N/A

## 72

TABLE 20-continued

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME00902 events containing Ceres Clone 36412					
	Event -01	Event -03	Event -04	Event -05	Control
Oil content (% control) in T <sub>3</sub> seed	104 ± 1	98 ± 1	102 ± 1	103 ± 3	100 ± 4
p-value***	<0.01	0.04	0.09	0.12	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME00902. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from one event of ME00902 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 20, the oil content was increased to 104% in seed from event-01 compared to the oil content in control seed. The oil content in T<sub>3</sub> seed from one event of ME00902 was significantly decreased compared to the oil content in corresponding control seed. As presented in Table 20, the oil content was decreased to 98% in seed from event-03 compared to the oil content in control seed.

## Example 19

## Results for ME00914 Events

T<sub>2</sub> and T<sub>3</sub> seed from five events and four events, respectively, of ME00914 containing Ceres Clone 8161 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from four events of ME00914 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME00914. As presented in Table 21, the oil content was increased to 121%, 122%, 125%, and 129% in seed from events-01, -02, -03, and -05, respectively, compared to the population mean.

TABLE 21

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME00914 events containing Ceres Clone 8161						
	Event -01	Event -02	Event -03	Event -04	Event -05	Control
Oil content (% control) in T <sub>2</sub> seed	121	122	125	116	129	100 ± 10*
p-value**	0.04	0.03	0.02	0.06	0.01	N/A
Oil content (% control) in T <sub>3</sub> seed	101 ± 0	98 ± 1	98 ± 2	No data	107 ± 1	100 ± 4
p-value***	0.46	<0.01	0.05	No data	<0.01	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME00914. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from one event of ME00914 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 21, the oil content was increased to 107% in seed from event-05 compared to the oil content in control seed. The oil content in T<sub>3</sub> seed from two events of ME00914 was significantly decreased compared to the oil content in corresponding control seed. As presented in Table 21, the oil content was decreased to 98% in seed from events-02 and -03 compared to the oil content in control seed.

## Results for ME01754 Events

T<sub>2</sub> and T<sub>3</sub> seed from five events and four events, respectively, of ME01754 containing Ceres Clone 368 was analyzed for oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from two events of ME01754 was significantly increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME01754. As presented in Table 22, the oil content was increased to 124% and 131% in seed from events-02 and -04, respectively, compared to the population mean.

TABLE 22

Oil content (% control) in T <sub>2</sub> and T <sub>3</sub> seed from ME01754 events containing Ceres Clone 368						
	Event - 01	Event - 02	Event - 03	Event - 04	Event - 05	Control
Oil content (% control) in T <sub>2</sub> seed	100	124	113	131	89	100 ± 9*
p-value**	0.18	0.01	0.08	<0.01	0.10	N/A
Oil content (% control) in T <sub>3</sub> seed	97 ± 2	103 ± 2	109 ± 3	96 ± 1	No data	100 ± 4
p-value***	0.11	0.02	0.10	<0.01	No data	N/A

\*Population mean of the oil content in seed from transgenic lines planted within 30 days of ME01754. Variation is presented as the standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

\*\*\*The p-values for T<sub>3</sub> seed were calculated using a Student's t-test.

The oil content in T<sub>3</sub> seed from one event of ME01754 was significantly increased compared to the oil content in corresponding control seed. As presented in Table 22, the oil content was increased to 103% in seed from event-02 compared to the oil content in control seed. The oil content in T<sub>3</sub> seed from one event of ME01754 was significantly decreased compared to the oil content in corresponding control seed. As presented in Table 22, the oil content was decreased to 96% in seed from event-04 compared to the oil content in control seed.

## Example 21

## Determination of Functional Homolog and/or Ortholog Sequences

A subject sequence was considered a functional homolog or ortholog of a query sequence if the subject and query sequences encoded proteins having a similar function and/or activity. A process known as Reciprocal BLAST (Rivera et al., *Proc. Natl. Acad. Sci. USA*, 95:6239-6244 (1998)) was used to identify potential functional homolog and/or ortholog sequences from databases consisting of all available public and proprietary peptide sequences, including NR from NCBI and peptide translations from Ceres clones.

Before starting a Reciprocal BLAST process, a specific query polypeptide was searched against all peptides from its source species using BLAST in order to identify polypeptides having BLAST sequence identity of 80% or greater to the query polypeptide and an alignment length of 85% or greater along the shorter sequence in the alignment. The query polypeptide and any of the aforementioned identified polypeptides were designated as a cluster.

The BLASTP version 2.0 program from Washington University at Saint Louis, Mo., USA was used to determine BLAST sequence identity and E-value. The BLASTP version 2.0 program includes the following parameters: 1) an E-value cutoff of 1.0e-5; 2) a word size of 5; and 3) the -postsw option. The BLAST sequence identity was calculated based on the alignment of the first BLAST HSP (High-scoring Segment Pairs) of the identified potential functional homolog and/or ortholog sequence with a specific query polypeptide. The number of identically matched residues in the BLAST HSP alignment was divided by the HSP length, and then multiplied by 100 to get the BLAST sequence identity. The HSP length typically included gaps in the alignment, but in some cases gaps were excluded.

The main Reciprocal BLAST process consists of two rounds of BLAST searches; forward search and reverse search. In the forward search step, a query polypeptide sequence, "polypeptide A," from source species SA was BLASTed against all protein sequences from a species of interest. Top hits were determined using an E-value cutoff of 10<sup>-5</sup> and a sequence identity cutoff of 35%. Among the top hits, the sequence having the lowest E-value was designated as the best hit, and considered a potential functional homolog or ortholog. Any other top hit that had a sequence identity of 80% or greater to the best hit or to the original query polypeptide was considered a potential functional homolog or ortholog as well. This process was repeated for all species of interest.

In the reverse search round, the top hits identified in the forward search from all species were BLASTed against all protein sequences from the source species SA. A top hit from the forward search that returned a polypeptide from the aforementioned cluster as its best hit was also considered as a potential functional homolog or ortholog.

Functional homologs and/or orthologs were identified by manual inspection of potential functional homolog and/or ortholog sequences. Representative functional homologs and/or orthologs for SEQ ID NO:81, SEQ ID NO:94, SEQ ID NO:111, SEQ ID NO:136, SEQ ID NO:152, SEQ ID NO:159, SEQ ID NO:171, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:193, SEQ ID NO:332, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346, SEQ ID NO:359, SEQ ID NO:374, and SEQ ID NO:398 are shown in FIGS. 1-13, respectively. The percent identities of functional homologs and/or orthologs to SEQ ID NO:81, SEQ ID NO:94, SEQ ID NO:111, SEQ ID NO:136, SEQ ID NO:152, SEQ ID NO:159, SEQ ID NO:171, SEQ ID NO:176, SEQ ID NO:178, SEQ ID NO:193, SEQ ID NO:332, SEQ ID NO:342, SEQ ID NO:344, SEQ ID NO:346, SEQ ID NO:359, SEQ ID NO:374, and SEQ ID NO:398 are shown below in Tables 23-34, respectively. The BLAST sequence identities and E-values given in Tables 23-34 were taken from the forward search round of the Reciprocal BLAST process.

TABLE 23

Percent identity to Ceres Clone 41573 (SEQ ID NO: 81)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres CLONE ID no. 1560908	<i>Zea mays</i>	82	54.9	2.49E-82
Public GI no. 2739168	<i>Nicotiana tabacum</i>	83	54.6	8.99E-85
Ceres CLONE ID no. 399052	<i>Zea mays</i>	84	54.5	1.19E-82

75

TABLE 23-continued

Percent identity to Ceres Clone 41573 (SEQ ID NO: 81)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres CLONE ID no. 1314177	<i>Triticum aestivum</i>	85	54.2	1.49E-82
Public GI no. 15824567	<i>Nicotiana tabacum</i>	86	53.9	2.09E-85
Ceres CLONE ID no. 1371577	<i>Glycine max</i>	87	53.2	2.90E-79
Public GI no. 15824565	<i>Nicotiana tabacum</i>	88	53.2	1.90E-84
Public GI no. 50920801	<i>Oryza sativa</i> subsp. <i>japonica</i>	89	53.2	1.80E-77
Public GI no. 50909807	<i>Oryza sativa</i> subsp. <i>japonica</i>	90	51.8	7.20E-76
Ceres CLONE ID no. 639223	<i>Triticum aestivum</i>	91	51.2	1.20E-75
Public GI no. 37531218	<i>Oryza sativa</i>	92	51.2	1.80E-72
Ceres CLONE ID no. 1476735	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	502	63.8	2.00E-96

TABLE 24

Percent identity to Ceres Clone 25429 (SEQ ID NO: 94)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres Annot: 1488311_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	96	75	1.70E-35
Ceres CLONE ID no. 953928	<i>Brassica napus</i>	97	74.7	1.49E-34
Ceres CLONE ID no. 524682	<i>Glycine max</i>	98	74.5	3.60E-33
Ceres CLONE ID no. 1609735	<i>Parthenium argentatum</i>	99	73.4	1.29E-28
Ceres CLONE ID no. 949174	<i>Brassica napus</i>	100	73.4	2.40E-36
Ceres CLONE ID no. 1299820	<i>Brassica napus</i>	101	72.5	3.90E-36
Public GI no. 42565379	<i>Hyacinthus orientalis</i>	102	72.5	1.30E-35
Ceres CLONE ID no. 426736	<i>Zea mays</i>	103	71.6	2.80E-35
Ceres CLONE ID no. 1094375	<i>Brassica napus</i>	104	71.4	1.39E-33
Ceres CLONE ID no. 691062	<i>Glycine max</i>	105	71.1	2.89E-33
Public GI no. 47026878	<i>Hyacinthus orientalis</i>	106	70.7	2.49E-34
Public GI no. 24473796	<i>Prunus dulcis</i>	107	69.9	8.20E-36
Ceres Annot: 1465437_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	109	67.2	3.39E-30
Ceres CLONE ID no. 1798334	<i>Gossypium hirsutum</i>	503	76.7	7.99E-38
Ceres CLONE ID no. 1886478	<i>Gossypium hirsutum</i>	504	76.3	9.50E-35
Ceres CLONE ID no. 1727128	<i>Musa acuminata</i>	505	72.9	6.89E-32
Public GI no. 730583	<i>Parthenium argentatum</i>	506	66	3.30E-32

76

TABLE 25

Percent identity to Ceres Clone 5750 (SEQ ID NO: 111)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Public GI no. 42362268	<i>Glycine max</i>	112	96.5	1.20E-41
Ceres CLONE ID no. 709027	<i>Glycine max</i>	113	96.5	1.20E-41
Ceres CLONE ID no. 853298	<i>Glycine max</i>	114	96.5	1.20E-41
Ceres CLONE ID no. 1417425	<i>Glycine max</i>	115	96.5	1.20E-41
Public GI no. 27435806	<i>Populus tremula</i> x <i>Populus tremuloides</i>	116	94.8	9.09E-37
Ceres Annot: 1481954_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	118	94.8	9.09E-37
Public GI no. 45935118	<i>Ipomoea trifida</i>	119	94.2	5.19E-41
Ceres CLONE ID no. 1017141	<i>Triticum aestivum</i>	120	93.8	3.79E-38
Ceres CLONE ID no. 1448636	<i>Zea mays</i>	121	93.8	3.79E-38
Public GI no. 50919707	<i>Oryza sativa</i> subsp. <i>japonica</i>	122	93.8	3.79E-38
Public GI no. 40641585	<i>Triticum aestivum</i>	123	93.8	3.79E-38
Public GI no. 38566522	<i>Arabidopsis thaliana</i>	124	93.75	4.39E-37
Ceres CLONE ID no. 1338131	<i>Arabidopsis thaliana</i>	125	93.75	4.39E-37
Ceres CLONE ID no. 300692	<i>Zea mays</i>	126	90.8	9.79E-40
Ceres CLONE ID no. 947192	<i>Brassica napus</i>	127	90.8	3.30E-39
Ceres CLONE ID no. 1465004	<i>Zea mays</i>	128	90.8	8.80E-39
Ceres CLONE ID no. 1122958	<i>Brassica napus</i>	129	89.6	6.19E-38
Ceres CLONE ID no. 944737	<i>Brassica napus</i>	130	86.4	9.69E-33
Ceres CLONE ID no. 217797	<i>Zea mays</i>	131	85.1	5.80E-35
Ceres CLONE ID no. 520185	<i>Glycine max</i>	132	85.1	5.80E-35
Public GI no. 55978016	<i>Ostreococcus tauri</i>	133	83	2.09E-30
Ceres CLONE ID no. 1436585	<i>Zea mays</i>	134	82.7	2.49E-34
Ceres CLONE ID no. 1777369	<i>Panicum virgatum</i>	507	96.1	1.70E-37
Ceres CLONE ID no. 1744578	<i>Musa acuminata</i>	508	95	5.49E-39
Ceres CLONE ID no. 100008703	<i>Gossypium hirsutum</i>	509	94.2	1.80E-40
Ceres CLONE ID no. 1723582	<i>Musa acuminata</i>	510	93.8	1.10E-38

TABLE 26

Percent identity to Ceres Clone 218626 (SEQ ID NO: 136)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Public GI no. 30409136	<i>Oryza sativa</i> subsp. <i>indica</i>	137	91.9	3.90E-219
Public GI no. 50940751	<i>Oryza sativa</i> subsp. <i>japonica</i>	138	91	3.29E-222
Ceres CLONE ID no. 1571117	<i>Zea mays</i>	139	84.9	3.70E-207
Ceres CLONE ID no. 424395	<i>Zea mays</i>	140	83.9	2.50E-201
Ceres Annot: 1440705_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	142	81.8	1.70E-161
Ceres Annot: 1493584_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	144	81.6	4.19E-174
Ceres Annot: 1463076_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	146	79.2	1.80E-159
Ceres CLONE ID no. 1002421	<i>Arabidopsis thaliana</i>	147	79.1	9.99E-171
Ceres Annot: 1516369_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	149	78.6	1.20E-153
Public GI no. 30693666	<i>Arabidopsis thaliana</i>	150	78.3	3.40E-188
Ceres CLONE ID no. 1796001	<i>Panicum virgatum</i>	511	85.8	1.19E-208

25

TABLE 27

Percent identity to Ceres Clone 121021 (SEQ ID NO: 152)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres Annot: 1501628_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	154	57.6	1.79E-17
Ceres Annot: 1519046_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	156	52.8	7.69E-24
Ceres CLONE ID no. 1121512	<i>Glycine max</i>	157	51.7	9.09E-21

TABLE 28

Percent identity to Ceres Clone 158765 (SEQ ID NO: 159)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Public GI no. 32562996	<i>Arabidopsis thaliana</i>	160	97.9	5.19E-73
Public GI no. 5669656	<i>Lycopersicon esculentum</i>	161	66.1	5.00E-36
Ceres CLONE ID no. 754061	<i>Triticum aestivum</i>	162	65.3	6.99E-30
Ceres CLONE ID no. 1329861	<i>Triticum aestivum</i>	163	65.3	6.99E-30
Ceres CLONE ID no. 537752	<i>Glycine max</i>	164	65.1	1.10E-40
Ceres CLONE ID no. 1322549	<i>Triticum aestivum</i>	165	64.2	1.90E-29
Ceres CLONE ID no. 282892	<i>Zea mays</i>	166	62.2	4.90E-29

TABLE 28-continued

30

Percent identity to Ceres Clone 158765 (SEQ ID NO: 159)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres CLONE ID no. 284046	<i>Zea mays</i>	167	62.2	1.29E-28
Ceres CLONE ID no. 1388825	<i>Zea mays</i>	168	61.6	4.90E-29
Public GI no. 50925813	<i>Oryza sativa</i> subsp. <i>japonica</i>	169	50	1.60E-30
Ceres CLONE ID no. 1839717	<i>Gossypium hirsutum</i>	545	72.8	5.09E-36

40

TABLE 29

Percent identity to Ceres Clone 16403 (SEQ ID NO: 171)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres CLONE ID no. 611156	<i>Glycine max</i>	172	58.9	7.29E-67
Ceres Annot: 1464944_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	174	56.2	4.39E-60
Ceres CLONE ID no. 1728680	<i>Musa acuminata</i>	530	60.2	1.70E-58
Ceres CLONE ID no. 1807796	<i>Gossypium hirsutum</i>	531	59.1	8.00E-70
Ceres CLONE ID no. 1771837	<i>Panicum virgatum</i>	532	52.3	4.69E-56
Ceres CLONE ID no. 1773482	<i>Panicum virgatum</i>	533	51.4	4.69E-56

45

TABLE 30

Percent identity to Ceres Clone 28635 (SEQ ID NO: 178)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Public GI no. 2463569	<i>Glycine max</i>	179	80.6	1.80E-180
Ceres	<i>Populus balsamifera</i>	181	79.9	6.20E-157

TABLE 30-continued

Percent identity to Ceres Clone 28635 (SEQ ID NO: 178)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Annot: 1514021_PRT	subsp. <i>trichocarpa</i>			
Public GI no. 55710094	<i>Centella asiatica</i>	182	79.8	2.49E-176
Ceres	<i>Populus balsamifera</i>	184	79.7	2.90E-166
Annot: 1503464_PRT	subsp. <i>trichocarpa</i>			
Public GI no. 75859951	<i>Panax notoginseng</i>	185	78.8	2.19E-177
Public GI no. 1449163	<i>Glycyrrhiza glabra</i>	186	78.8	5.19E-176
Public GI no. 1449165	<i>Glycyrrhiza glabra</i>	187	78.1	3.69E-175
Public GI no. 28208268	<i>Lotus japonicus</i>	188	78.1	1.60E-172
Public GI no. 41224629	<i>Panax ginseng</i>	189	77.8	3.69E-175
Public GI no. 27475614	<i>Medicago truncatula</i>	190	77.8	2.10E-172
Public GI no. 38426486	<i>Artemisia annua</i>	191	77	8.10E-171
Ceres CLONE ID no. 1920025	<i>Gossypium hirsutum</i>	534	79.2	9.80E-175
Public GI no. 110293133	<i>Polygala tenuifolia</i>	535	78.6	2.29E-175
Public GI no. 5360655	<i>Solanum tuberosum</i>	536	76.0	1.5E-171
Public GI no. 4426953	<i>Capsicum annuum</i>	537	75.7	2.2E-170
Public GI no. 1552717	<i>Nicotiana tabacum</i>	538	75.5	1.9E-171
Public GI no. 66393825	<i>Bupleurum falcatum</i>	539	75.9	1.69E-170
Public GI no. 1706774	<i>Nicotiana benthamiana</i>	540	74.8	2.5E-169
Ceres CLONE ID no. 1749989	<i>Panicum virgatum</i>	541	69.8	3.20E-153
Public GI no. 115456049	<i>Oryza sativa</i> subsp. <i>japonica</i>	542	68.6	1.60E-149
Public GI no. 2463567	<i>Zea mays</i>	543	68.3	3.79E-150
Ceres CLONE ID no. 706088	<i>Triticum aestivum</i>	544	65.1	5.9E-145

TABLE 31

Percent identity to Ceres Clone 35698 (SEQ ID NO: 193)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres CLONE ID no. 1346445	<i>Arabidopsis thaliana</i>	194	98.6	1.90E-36
Public GI no. 441457	<i>Lycopersicon esculentum</i>	195	98	8.20E-52
Public GI no. 19347859	<i>Arabidopsis thaliana</i>	196	98	3.49E-51
Ceres	<i>Populus balsamifera</i>	198	97	2.80E-51
Annot: 1483290_PRT	subsp. <i>trichocarpa</i>			
Public GI no. 40287554	<i>Capsicum annuum</i>	199	97	1.29E-51
Public GI no. 21553796	<i>Arabidopsis thaliana</i>	200	97	2.20E-51
Public GI no. 28569271	<i>Gossypium thurberi</i>	201	97	2.20E-51
Public GI no. 28569265	<i>Gossypium hirsutum</i>	202	97	2.20E-51
Public GI no. 66354420	<i>Arabidopsis thaliana</i>	203	97	2.20E-51
Public GI no. 21280893	<i>Arabidopsis thaliana</i>	204	97	2.20E-51
Public GI no. 21554343	<i>Arabidopsis thaliana</i>	205	97	2.20E-51
Public GI no. 22597164	<i>Glycine max</i>	206	97	4.50E-51
Ceres CLONE ID no. 617835	<i>Triticum aestivum</i>	207	96	9.40E-51
Public GI no. 30693871	<i>Arabidopsis thaliana</i>	208	96	8.40E-50
Public GI no. 5762457	<i>Mesembryanthemum crystallinum</i>	209	96	4.50E-51
Public GI no. 464981	<i>Lycopersicon esculentum</i>	210	96	4.50E-51
Public GI no. 456568	<i>Pisum sativum</i>	211	96	4.50E-51
Public GI no. 77416935	<i>Solanum tuberosum</i>	212	96	3.49E-51
Public GI no. 28569267	<i>Gossypium arboreum</i>	213	96	3.49E-51
Public GI no. 28569261	<i>Gossypium hirsutum</i>	214	96	3.49E-51
Ceres CLONE ID no. 39130	<i>Arabidopsis thaliana</i>	215	95.8	9.99E-36
Ceres CLONE ID no. 16865	<i>Arabidopsis thaliana</i>	216	95.8	9.99E-36
Ceres CLONE ID no. 575067	<i>Glycine max</i>	217	95.8	1.30E-35
Ceres	<i>Populus balsamifera</i>	219	95.8	1.30E-35
Annot: 1467392_PRT	subsp. <i>trichocarpa</i>			
Ceres CLONE ID no. 25162	<i>Arabidopsis thaliana</i>	220	95.8	9.99E-36
Ceres	<i>Populus balsamifera</i>	222	95.8	8.20E-36
Annot: 1529647_PRT	subsp. <i>trichocarpa</i>			
Ceres CLONE ID no. 1405728	<i>Zea mays</i>	223	95.8	8.20E-36

TABLE 31-continued

Percent identity to Ceres Clone 35698 (SEQ ID NO: 193)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Public GI no. 54288726	<i>Picea abies</i>	224	95.3	9.99E-31
Public GI no. 50906823	<i>Oryza sativa</i> subsp. japonica	225	95	1.20E-50
Public GI no. 83306206	<i>Oryza sativa</i> subsp. indica	226	95	3.19E-50
Public GI no. 20152203	<i>Oryza sativa</i> subsp. japonica	227	95	1.89E-50
Public GI no. 40287568	<i>Capsicum annuum</i>	228	94.5	2.20E-35
Public GI no. 50929483	<i>Oryza sativa</i> subsp. japonica	229	94.1	4.00E-50
Ceres Annot: 1450556_PRT	<i>Populus balsamifera</i> subsp. trichocarpa	231	93.1	4.50E-35
Ceres CLONE ID no. 1031152	<i>Triticum aestivum</i>	232	93.1	8.40E-50
Public GI no. 52548244	<i>Triticum aestivum</i>	233	93.1	8.40E-50
Public GI no. 82621144	<i>Solanum tuberosum</i>	234	92.8	7.60E-33
Ceres CLONE ID no. 878043	<i>Triticum aestivum</i>	235	92.1	2.80E-49
Public GI no. 34909292	<i>Oryza sativa</i> subsp. japonica	236	92.1	2.80E-49
Public GI no. 2668744	<i>Zea mays</i>	237	92.1	2.80E-49
Ceres CLONE ID no. 511132	<i>Glycine max</i>	238	91.7	9.40E-35
Ceres Annot: 1533930_PRT	<i>Populus balsamifera</i> subsp. trichocarpa	240	91.7	1.19E-34
Ceres Annot: 1495171_PRT	<i>Populus balsamifera</i> subsp. trichocarpa	242	91.1	9.99E-38
Public GI no. 20086317	<i>Oryza sativa</i>	243	90.5	3.19E-50
Ceres CLONE ID no. 10022	<i>Arabidopsis thaliana</i>	244	90.4	4.09E-34
Ceres CLONE ID no. 12547	<i>Arabidopsis thaliana</i>	245	90.4	4.09E-34
Ceres CLONE ID no. 27679	<i>Arabidopsis thaliana</i>	246	90.4	4.09E-34
Public GI no. 20259611	<i>Arabidopsis thaliana</i>	247	90.1	9.60E-49
Public GI no. 30025160	<i>Hordeum vulgare</i>	248	90.1	5.30E-48
Public GI no. 50904839	<i>Oryza sativa</i> subsp. japonica	249	89.2	9.60E-49
Ceres CLONE ID no. 1063753	<i>Zea mays</i>	250	89.2	8.69E-48
Ceres Annot: 1533218_PRT	<i>Populus balsamifera</i> subsp. trichocarpa	252	89	1.09E-33
Public GI no. 297880	<i>Arabidopsis thaliana</i>	253	88.8	2.89E-33
Ceres CLONE ID no. 1357060	<i>Zea mays</i>	254	88.2	1.09E-47
Public GI no. 54402104	<i>Arachis hypogaea</i>	255	87.2	2.89E-47
Public GI no. 50725323	<i>Oryza sativa</i> subsp. japonica	256	87.2	1.29E-46
Ceres CLONE ID no. 376667	<i>Zea mays</i>	257	87.2	9.89E-47
Ceres CLONE ID no. 256705	<i>Zea mays</i>	258	86.8	1.50E-41
Public GI no. 20259629	<i>Arabidopsis thaliana</i>	259	86.2	4.80E-47
Public GI no. 52851174	<i>Plantago major</i>	260	86.2	6.09E-47
Ceres CLONE ID no. 1061097	<i>Zea mays</i>	261	86.2	1.60E-46
Public GI no. 1373001	<i>Oryza sativa</i>	262	86.2	4.39E-44
Public GI no. 66354468	<i>Arabidopsis thaliana</i>	263	86.2	9.89E-47
Public GI no. 4100646	<i>Pinus resinosa</i>	264	82.3	2.40E-45
Ceres CLONE ID no. 1380019	<i>Zea mays</i>	265	98	2.80E-51
Ceres CLONE ID no. 1380019_T	Artificial Sequence	266	98.04	4.80E-56
Ceres CLONE ID no. 1346445_T	Artificial Sequence	267	98.63	3.30E-41
Public GI no. 441457_T	Artificial Sequence	268	98.04	1.40E-56
Public GI no. 19347859_T	Artificial Sequence	269	98.04	6.10E-56
Ceres Annot: 1483290_T	<i>Populus balsamifera</i> subsp. trichocarpa	270	97.06	4.80E-56
Public GI no. 40287554_T	Artificial Sequence	271	97.06	2.30E-56
Public GI no. 21553796_T	Artificial Sequence	272	97.06	3.70E-56
Public GI no. 28569271_T	Artificial Sequence	273	97.06	3.70E-56

TABLE 31-continued

Percent identity to Ceres Clone 35698 (SEQ ID NO: 193)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Public GI no. 28569265_T	Artificial Sequence	274	97.06	3.70E-56
Public GI no. 66354420_T	Artificial Sequence	275	97.06	3.70E-56
Public GI no. 21280893_T	Artificial Sequence	276	97.06	3.70E-56
Public GI no. 21554343_T	Artificial Sequence	277	97.06	3.70E-56
Public GI no. 22597164_T	Artificial Sequence	278	97.03	7.80E-56
Ceres CLONE ID no. 617835_T	Artificial Sequence	279	96.08	1.60E-55
Public GI no. 30693871_T	Artificial Sequence	280	96.08	1.50E-54
Public GI no. 5762457_T	Artificial Sequence	281	96.08	7.80E-56
Public GI no. 464981_T	Artificial Sequence	282	96.08	7.80E-56
Public GI no. 456568_T	Artificial Sequence	283	96.08	7.80E-56
Public GI no. 77416935_T	Artificial Sequence	284	96.08	6.10E-56
Public GI no. 28569267_T	Artificial Sequence	285	96.08	6.10E-56
Public GI no. 28569261_T	Artificial Sequence	286	96.08	6.10E-56
Ceres CLONE ID no. 39130_T	Artificial Sequence	287	95.89	1.80E-40
Ceres CLONE ID no. 16865_T	Artificial Sequence	288	95.89	1.80E-40
Ceres CLONE ID no. 575067_T	Artificial Sequence	289	95.89	2.30E-40
Ceres Annot: 1467392_T	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	290	95.89	2.30E-40
Ceres CLONE ID no. 25162_T	Artificial Sequence	291	95.89	1.80E-40
Ceres Annot: 1529647_T	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	292	95.89	1.40E-40
Ceres CLONE ID no. 1405728_T	Artificial Sequence	293	95.89	1.40E-40
Public GI no. 54288726_T	Artificial Sequence	294	95.38	1.70E-35
Public GI no. 50906823_T	Artificial Sequence	295	95.1	2.10E-55
Public GI no. 83306206_T	Artificial Sequence	296	95.1	5.50E-55
Public GI no. 20152203_T	Artificial Sequence	297	95.1	3.40E-55
Public GI no. 40287568_T	Artificial Sequence	298	94.52	3.80E-40
Public GI no. 50929483_T	Artificial Sequence	299	94.12	7.00E-55
Ceres Annot: 1450556_T	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	300	93.15	7.80E-40
Ceres CLONE ID no. 1031152_T	Artificial Sequence	301	93.14	1.50E-54
Public GI no. 52548244_T	Artificial Sequence	302	93.14	1.50E-54
Public GI no. 82621144_T	Artificial Sequence	303	92.86	1.30E-37
Ceres CLONE ID no. 878043_T	Artificial Sequence	304	92.16	4.90E-54
Public GI no. 34909292_T	Artificial Sequence	305	92.16	4.90E-54
Public GI no. 2668744_T	Artificial Sequence	306	92.16	4.90E-54
Ceres CLONE ID no. 511132_T	Artificial Sequence	307	91.78	1.60E-39
Ceres Annot: 1533930_T	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	308	91.78	2.10E-39
Ceres Annot: 1495171_T	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	309	91.14	1.80E-42
Public GI no. 20086317_T	Artificial Sequence	310	94.12	7.00E-55
Ceres CLONE ID no. 10022_T	Artificial Sequence	311	90.41	7.00E-39
Ceres CLONE ID no. 12547_T	Artificial Sequence	312	90.41	7.00E-39



TABLE 31-continued

Percent identity to Ceres Clone 35698 (SEQ ID NO: 193)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres CLONE ID no. 27679_T	Artificial Sequence	313	90.41	7.00E-39
Public GI no. 20259611_T	Artificial Sequence	314	90.2	1.70E-53
Public GI no. 30025160_T	Artificial Sequence	315	90.2	9.20E-53
Public GI no. 50904839_T	Artificial Sequence	316	89.22	1.70E-53
Ceres CLONE ID no. 1063753_T	Artificial Sequence	317	89.22	1.50E-52
Ceres Annot: 1533218_T	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	318	89.04	1.90E-38
Public GI no. 297880_T	Artificial Sequence	319	88.89	4.90E-38
Ceres CLONE ID no. 1357060_T	Artificial Sequence	320	88.24	1.90E-52
Public GI no. 54402104_T	Artificial Sequence	321	87.25	5.10E-52
Public GI no. 50725323_T	Artificial Sequence	322	87.25	2.20E-51
Ceres CLONE ID no. 376667_T	Artificial Sequence	323	87.25	1.70E-51
Ceres CLONE ID no. 256705_T	Artificial Sequence	324	86.81	2.70E-46
Public GI no. 20259629_T	Artificial Sequence	325	86.27	8.30E-52
Public GI no. 52851174_T	Artificial Sequence	326	86.27	1.10E-51
Ceres CLONE ID no. 1061097_T	Artificial Sequence	327	86.27	2.80E-51
Public GI no. 1373001_T	Artificial Sequence	328	86.27	7.70E-49
Public GI no. 66354468_T	Artificial Sequence	329	86.27	1.70E-51
Public GI no. 4100646_T	Artificial Sequence	330	82.35	4.10E-50

TABLE 32

Percent identity to Ceres Clone 36412 (SEQ ID NO: 332)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Public GI no. 3152583	<i>Arabidopsis thaliana</i>	333	98.5	2.00E-176
Ceres Annot: 1467033_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	335	60.2	1.20E-68
Ceres Annot: 1536919_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	337	60.1	7.90E-70
Ceres CLONE ID no. 1641329	<i>Glycine max</i>	338	52.7	1.90E-61
Ceres CLONE ID no. 1650419	<i>Glycine max</i>	339	52	1.39E-65
Ceres CLONE ID no. 597699	<i>Glycine max</i>	340	51	1.09E-63

TABLE 33

Percent identity to Ceres Clone 4829 (SEQ ID NO: 346)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres CLONE ID no. 24885	<i>Arabidopsis thaliana</i>	347	87.2	8.89E-62

35

TABLE 33-continued

Percent identity to Ceres Clone 4829 (SEQ ID NO: 346)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres CLONE ID no. 27878	<i>Arabidopsis thaliana</i>	348	84.2	1.19E-57
Ceres Annot: 1485102_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	350	81.2	1.50E-52
Ceres CLONE ID no. 1646533	<i>Glycine max</i>	351	72.8	2.30E-54
Public GI no. 29371519	<i>Oryza sativa</i> subsp. <i>japonica</i>	352	63.4	5.99E-40
Public GI no. 38347602	<i>Oryza sativa</i> subsp. <i>japonica</i>	353	63.4	5.99E-40
Public GI no. 45935148	<i>Ipomoea trifida</i>	354	60.8	1.29E-39
Ceres CLONE ID no. 359934	<i>Zea mays</i>	355	60.4	3.79E-38
Ceres CLONE ID no. 294598	<i>Zea mays</i>	356	58.8	5.99E-40
Ceres CLONE ID no. 839270	<i>Triticum aestivum</i>	357	55.1	3.10E-36
Ceres CLONE ID no. 1836904	<i>Gossypium hirsutum</i>	525	78.6	1.40E-54
Ceres CLONE ID no. 1932013	<i>Gossypium hirsutum</i>	526	77.3	7.99E-54
Ceres CLONE ID no. 1768109	<i>Panicum virgatum</i>	527	58.1	7.79E-40

65

87

TABLE 34

Percent identity to Ceres Clone 5426 (SEQ ID NO: 359)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres CLONE ID no. 1123542	<i>Brassica napus</i>	360	90.3	3.99E-59
Ceres CLONE ID no. 9083	<i>Arabidopsis thaliana</i>	361	90.2	4.80E-63
Public GI no. 79322493	<i>Arabidopsis thaliana</i>	362	90.2	4.80E-63
Ceres CLONE ID no. 265408	<i>Arabidopsis thaliana</i>	363	90.2	4.80E-63
Ceres CLONE ID no. 32164	<i>Arabidopsis thaliana</i>	364	90.2	4.80E-63
Ceres CLONE ID no. 1068047	<i>Brassica napus</i>	365	88.1	2.39E-61
Ceres CLONE ID no. 965035	<i>Brassica napus</i>	366	87.4	6.30E-61
Ceres Annot: 1499194_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	368	76.7	1.00E-53
Ceres Annot: 1439584_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	370	75.5	3.40E-53
Ceres CLONE ID no. 557065	<i>Glycine max</i>	371	73.4	5.60E-53
Ceres CLONE ID no. 465060	<i>Glycine max</i>	372	72.5	1.29E-51
Ceres CLONE ID no. 1458107	<i>Zea mays</i>	512	87.4	2.99E-61

88

TABLE 34-continued

Percent identity to Ceres Clone 5426 (SEQ ID NO: 359)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Ceres CLONE ID no. 1932511	<i>Gossypium hirsutum</i>	513	76.2	1.10E-54
Ceres CLONE ID no. 1850967	<i>Gossypium hirsutum</i>	514	75.5	7.99E-54
Ceres CLONE ID no. 1835707	<i>Gossypium hirsutum</i>	515	75.5	1.00E-53
Ceres CLONE ID no. 1727213	<i>Musa acuminata</i>	516	72.7	5.19E-50
Ceres CLONE ID no. 1767577	<i>Panicum virgatum</i>	517	71.3	9.69E-49
Ceres CLONE ID no. 1712104	<i>Musa acuminata</i>	518	71.3	8.50E-50
Ceres CLONE ID no. 1778377	<i>Panicum virgatum</i>	519	71.1	1.20E-48
Public GI no. 76573317	<i>Solanum tuberosum</i>	520	70.6	4.10E-50
Ceres CLONE ID no. 1787980	<i>Panicum virgatum</i>	521	69.9	1.29E-46
Public GI no. 115465181	<i>Oryza sativa</i> subsp. <i>japonica</i>	522	69.7	6.89E-48
Public GI no. 48716267	<i>Oryza saliva</i> subsp. <i>japonica</i>	523	69.7	8.69E-48
Ceres CLONE ID no. 575833	<i>Triticum aestivum</i>	524	66.1	3.39E-46

TABLE 35

Percent identity to Ceres Clone 7894 (SEQ ID NO: 374)				
Designation	Species	SEQ ID NO:	% Identity	e-value
Public GI no. 18091781	<i>Brassica oleracea</i>	375	94.6	6.89E-261
Public GI no. 468562	<i>Ricinus communis</i>	376	75.3	2.39E-194
Ceres Annot: 1479767_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	378	74.7	4.30E-188
Public GI no. 7649151	<i>Euphorbia esula</i>	379	74.3	9.99E-194
Ceres Annot: 1486712_PRT	<i>Populus balsamifera</i> subsp. <i>trichocarpa</i>	381	72.5	3.70E-182
Public GI no. 33620334	<i>Glycine max</i>	382	71.9	3.29E-181
Public GI no. 439294	<i>Solanum tuberosum</i>	383	71.6	4.39E-179
Public GI no. 77153413	<i>Populus tremula</i> x <i>Populus tremuloides</i>	384	71.6	8.90E-181
Public GI no. 5230818	<i>Pisum sativum</i>	385	71.1	1.40E-180
Public GI no. 17447420	<i>Alonsoa meridionalis</i>	386	70.9	1.29E-179
Public GI no. 1935019	<i>Vicia faba</i>	387	70.8	4.39E-179
Public GI no. 51863031	<i>Juglans regia</i>	388	69.6	8.79E-174
Public GI no. 575351	<i>Nicotiana tabacum</i>	389	69.3	4.39E-179
Public GI no. 68161544	<i>Vitis vinifera</i>	390	68.7	7.39E-177
Public GI no. 6434833	<i>Vitis vinifera</i>	391	68.5	1.49E-176
Public GI no. 21319	<i>Spinacia oleracea</i>	392	68.2	9.99E-173
Public GI no. 5823000	<i>Beta vulgaris</i>	393	67.7	2.69E-170
Public GI no. 633172	<i>Beta vulgaris</i> subsp. <i>vulgaris</i>	394	67.3	8.99E-172
Public GI no. 6120115	<i>Asarina barclaiana</i>	395	67.2	5.50E-172
Public GI no. 415988	<i>Plantago major</i>	396	66.2	7.39E-177
Public GI no. 116008246	<i>Hevea brasiliensis</i>	528	73	9.09E-188
Ceres CLONE ID no. 1925996	<i>Gossypium hirsutum</i>	529	66.8	4.09E-176

## Transgenic Plants Containing Homologs and/or Orthologs

Cloned sequences of some of the functional homologs and/or orthologs of protein-modulating polypeptides that were identified as outlined in Example 20 were used to make transgenic plants.

Ceres Clone 39130 (SEQ ID NO:440) is a cDNA clone isolated from *Arabidopsis* that encodes a functional homologue of SEQ ID NO:193, and is predicted to encode a 119 amino acid transcription ubiquitin-conjugating enzyme polypeptide. Ceres Clone 424395 (SEQ ID NO:427) is a cDNA clone isolated from *Zea mays* that encodes a functional homologue of SEQ ID NO:136, and is predicted to encode a 446 amino acid tryptophan/tyrosine permease family polypeptide. Ceres Clone 24885 (SEQ ID NO:460) is a cDNA clone isolated from *Arabidopsis* that encodes a functional homologue of SEQ ID NO:346, and is predicted to encode a 156 amino acid polypeptide having a DUF662 domain. Ceres Clone 300692 (SEQ ID NO:418) is a cDNA clone isolated from *Zea mays* that encodes a functional homologue of SEQ ID NO:111, and is predicted to encode an 87 amino acid cyclin-dependent kinase regulatory subunit polypeptide. Ceres Clone 944737 (SEQ ID NO:422) is a cDNA clone isolated from *Brassica napus* that encodes a functional homologue of SEQ ID NO:111, and is predicted to encode an 81 amino acid cyclin-dependent kinase regulatory subunit polypeptide.

A construct was made with the CRS 311 vector that contained Ceres Clone 39130 operably linked to the 32449 promoter. Constructs were made with the CRS 338 vector that contained Ceres Clone 424395, Ceres Clone 24885, Ceres Clone 300692, or Ceres Clone 944737, each operably linked to a CaMV 35S promoter. Wild-type *Arabidopsis thaliana* ecotype Wassilewskija (Ws) plants were transformed separately with each construct as described in Example 1.

Transgenic *Arabidopsis* lines containing Ceres Clone 39130, Ceres Clone 424395, Ceres Clone 24885, Ceres Clone 300692, or Ceres Clone 944737 were designated ME00177, ME06339, ME07409, ME08044, or ME08488, respectively. The presence of each vector containing a Ceres clone described above in the respective transgenic *Arabidopsis* line transformed with the vector was confirmed by Finale™ resistance, polymerase chain reaction (PCR) amplification from green leaf tissue extract, and/or sequencing of PCR products. As controls, wild-type *Arabidopsis* ecotype Ws plants were transformed with the empty vector CRS 338 or the empty vector CRS 311.

## Example 22

## Results for Transgenic Plants Containing Homologs and/or Orthologs

T<sub>2</sub> seed from three events of ME00177 containing Ceres Clone 39130 was analyzed for total oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from three events of ME00177 was increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME00177. As presented in Table 36, the oil content was increased to 115%, 109%, and 122% in seed from events-01, -02, and -03, respectively, compared to the population mean.

TABLE 36

Oil content (% control) in T <sub>2</sub> seed from ME00177 events containing Ceres Clone 39130				
	Event -01	Event -02	Event -03	Control
Protein content (% control) in T <sub>2</sub> seed	115	109	122	100 ± 7*
p-value**	0.04	0.15	<0.01	N/A

\*Population mean of the protein content in seed from transgenic lines planted within 30 days of ME00177. Variation is presented as standard error of the mean.  
\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

T<sub>2</sub> seed from three events of ME06339 containing Ceres Clone 424395 was analyzed for total oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from three events of ME06339 was increased compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME06339. As presented in Table 37, the protein content was increased to 110%, 103%, and 110% in seed from events-02, -03, and -04, respectively, compared to the population mean.

TABLE 37

Oil content (% control) in T <sub>2</sub> seed from ME06339 events containing Ceres Clone 424395				
	Event -02	Event -03	Event -04	Control
Protein content (% control) in T <sub>2</sub> seed	110	103	110	100 ± 12*
p-value	0.06	0.401	0.068	N/A

\*Population mean of the protein content in seed from transgenic lines planted within 30 days of ME06339. Variation is presented as standard error of the mean.

T<sub>2</sub> seed from four events of ME07409 containing Ceres Clone 24885 was analyzed for total oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from four events of ME07409 was modulated compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME07409. As presented in Table 38, the oil content was increased to 107%, 103%, and 108% in seed from events-02, -03, and -04, respectively, compared to the population mean, while the oil content was decreased to 98% of the population mean in event-06.

TABLE 38

Oil content (% control) in T <sub>2</sub> seed from ME07409 events containing Ceres Clone 24885					
	Event -02	Event -03	Event -04	Event -06	Control
Protein content (% control) in T <sub>2</sub> seed	107	103	108	108	100 ± 13*
p-value**	0.06	0.224	0.04	0.551	N/A

\*Population mean of the protein content in seed from transgenic lines planted within 30 days of ME07409. Variation is presented as standard error of the mean.  
\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

T<sub>2</sub> seed from four events of ME08044 containing Ceres Clone 300692 was analyzed for total oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from four events of ME08044 was modulated compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of

91

ME08044. As presented in Table 39, the protein content was increased to 108% and 105% in seed from events-01 and -02, respectively, compared to the population mean, while the oil content was decreased to 95% and 88% of the population mean in events-03 and -04, respectively.

TABLE 39

Oil content (% control) in T <sub>2</sub> seed from ME08044 events containing Ceres Clone 300692					
	Event -01	Event -02	Event -03	Event -04	Control
Protein content (% control) in T <sub>2</sub> seed	108	105	95	88	100 ± 15*
p-value**	0.094	0.215	0.112	<0.01	N/A

\*Population mean of the protein content in seed from transgenic lines planted within 30 days of ME08044. Variation is presented as standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

T<sub>2</sub> seed from five events of ME08488 containing Ceres Clone 944737 was analyzed for total oil content using FT-NIR spectroscopy as described in Example 2.

The oil content in T<sub>2</sub> seed from five events of ME08488 was modulated compared to the mean oil content in seed from transgenic *Arabidopsis* lines planted within 30 days of ME08488. As presented in Table 40, the protein content was increased to 103% and 107% in seed from events-02 and -04, respectively, compared to the population mean, while the oil

92

content was decreased to 97%, 84%, and 90% of the population mean in events-01, -03, and -05, respectively.

TABLE 40

Oil content (% control) in T <sub>2</sub> seed from ME08488 events containing Ceres Clone 944737						
	Event -01	Event -02	Event -03	Event -04	Event -05	Control
Protein content (% control) in T <sub>2</sub> seed	97	103	84	107	90	100 ± 18*
p-value**	0.48	0.248	0.104	0.148	0.746	N/A

\*Population mean of the protein content in seed from transgenic lines planted within 30 days of ME08488. Variation is presented as standard error of the mean.

\*\*The p-values for T<sub>2</sub> seed were calculated using z-scores.

Transgenic plants containing cloned sequences of some of the other functional homologs and/or orthologs of Example 21 were analyzed for total oil content in seeds by FT-NIR spectroscopy. The results were inconclusive.

#### Other Embodiments

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

#### SEQUENCE LISTING

The patent contains a lengthy "Sequence Listing" section. A copy of the "Sequence Listing" is available in electronic form from the USPTO web site (<http://seqdata.uspto.gov/?pageRequest=docDetail&DocID=US08222482B2>). An electronic copy of the "Sequence Listing" will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

What is claimed is:

1. A method of modulating the level of oil in a plant, said method comprising introducing into a plant cell an isolated nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to the amino acid sequence set forth in SEQ ID NO: 136, said polypeptide having tryptophan/tyrosine permease activity, wherein a tissue of a plant produced from said plant cell has a difference in the level of oil as compared to the corresponding level in tissue of a control plant that does not comprise said nucleic acid.

2. The method of claim 1, wherein said sequence identity is 85 percent or greater.

3. The method of claim 2, wherein said sequence identity is 90 percent or greater.

4. The method of claim 2, wherein said sequence identity is 95 percent or greater.

5. The method of claim 1, wherein said nucleotide sequence encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:136.

6. The method of claim 1, wherein said difference is an increase in the level of oil.

7. The method of claim 1, wherein said isolated nucleic acid is operably linked to a regulatory region.

8. The method of claim 7, wherein said regulatory region is a tissue-preferential regulatory region.

9. The method of claim 8, wherein said tissue-preferential regulatory region is a promoter.

10. The method of claim 7, wherein said regulatory region is a broadly expressing promoter.

11. The method of claim 1, wherein said plant is a dicot.

12. The method of claim 11, wherein said plant is a member of the genus *Anacardium*, *Arachis*, *Azadirachta*, *Brassica*, *Cannabis*, *Carthamus*, *Corylus*, *Crambe*, *Cucurbita*, *Glycine*, *Gossypium*, *Helianthus*, *Jatropha*, *Juglans*, *Linum*, *Olea*, *Papaver*, *Persea*, *Prunus*, *Ricinus*, *Sesamum*, *Simmondsia*, or *Vitis*.

13. The method of claim 1, wherein said plant is a monocot.

14. The method of claim 13, wherein said plant is a member of the genus *Cocos*, *Elaeis*, *Oryza*, or *Zea*.

15. The method of claim 1, wherein said tissue is seed tissue.

16. A method of producing a plant tissue, said method comprising growing a plant cell comprising an exogenous nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to the amino acid sequence set forth in SEQ ID NO: 136, said polypeptide having tryptophan/tyrosine permease activity,

wherein said tissue has a difference in the level of oil as compared to the corresponding level in tissue of a control plant that does not comprise said nucleic acid.

17. The method of claim 16, wherein said sequence identity is 85 percent or greater.

18. The method of claim 17, wherein said sequence identity is 90 percent or greater.

19. The method of claim 17, wherein said sequence identity is 95 percent or greater.

20. The method of claim 16, wherein said nucleotide sequence encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:136.

21. The method of claim 16, wherein said difference is an increase in the level of oil.

22. The method of claim 16, wherein said exogenous nucleic acid is operably linked to a regulatory region.

23. The method of claim 22, wherein said regulatory region is a tissue-preferential regulatory region.

24. The method of claim 23, wherein said tissue-preferential regulatory region is a promoter.

25. The method of claim 22, wherein said regulatory region is a broadly expressing promoter.

26. The method of claim 16, wherein said plant tissue is dicotyledonous.

27. The method of claim 26, wherein said plant tissue is a member of the genus *Anacardium*, *Arachis*, *Azadirachta*, *Brassica*, *Cannabis*, *Carthamus*, *Corylus*, *Crambe*, *Cucurbita*, *Glycine*, *Gossypium*, *Helianthus*, *Jatropha*, *Juglans*, *Linum*, *Olea*, *Papaver*, *Persea*, *Prunus*, *Ricinus*, *Sesamum*, *Simmondsia*, or *Vitis*.

28. The method of claim 16, wherein said plant tissue is monocotyledonous.

29. The method of claim 28, wherein said plant tissue is a member of the genus *Cocos*, *Elaeis*, *Oryza*, or *Zea*.

30. The method of claim 16, wherein said tissue is seed tissue.

31. A plant cell comprising an exogenous nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to the amino acid sequence set forth in SEQ ID NO:136, said polypeptide having tryptophan/tyrosine permease activity, wherein a tissue of a plant produced from said plant cell has a difference in the level of oil as compared to the corresponding level in tissue of a control plant that does not comprise said nucleic acid.

32. The plant cell of claim 31, wherein said sequence identity is 85 percent or greater.

33. The plant cell of claim 32, wherein said sequence identity is 90 percent or greater.

34. The plant cell of claim 32, wherein said sequence identity is 95 percent or greater.

35. The plant cell of claim 31, wherein said nucleotide sequence encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:136.

36. The plant cell of claim 31, wherein said difference is an increase in the level of oil.

37. The plant cell of claim 31, wherein said exogenous nucleic acid is operably linked to a regulatory region.

38. The plant cell of claim 37, wherein said regulatory region is a tissue-preferential regulatory region.

39. The plant cell of claim 38, wherein said tissue-preferential regulatory region is a promoter.

40. The plant cell of claim 37, wherein said regulatory region is a broadly expressing promoter.

41. The plant cell of claim 31, wherein said plant is a dicot.

42. The plant cell of claim 41, wherein said plant is a member of the genus *Anacardium*, *Arachis*, *Azadirachta*, *Brassica*, *Cannabis*, *Carthamus*, *Corylus*, *Crambe*, *Cucurbita*, *Glycine*, *Gossypium*, *Helianthus*, *Jatropha*, *Juglans*, *Linum*, *Olea*, *Papaver*, *Persea*, *Prunus*, *Ricinus*, *Sesamum*, *Simmondsia*, or *Vitis*.

43. The plant cell of claim 31, wherein said plant is a monocot.

44. The plant cell of claim 43, wherein said plant is a member of the genus *Cocos*, *Elaeis*, *Oryza*, or *Zea*.

45. The plant cell of claim 31, wherein said tissue is seed tissue.

46. A transgenic plant comprising the plant cell of claim 31.

47. Progeny of the plant of claim 46, wherein said progeny has a difference in the level of oil as compared to the level of oil in a corresponding control plant that does not comprise said exogenous nucleic acid.

48. Seed from a transgenic plant according to claim 46.

49. Vegetative tissue from a transgenic plant according to claim 46.

50. A food product comprising seed or vegetative tissue from a transgenic plant according to claim 46.

51. A feed product comprising seed or vegetative tissue from a transgenic plant according to claim 46.

52. A method of making oil, said method comprising extracting oil from the seed of claim 48.

53. A method of modulating the level of oil in a plant, said method comprising:

a) introducing into a plant cell an isolated nucleic acid comprising a nucleotide sequence encoding a polypeptide having 80 percent or greater sequence identity to the amino acid sequence set forth in SEQ ID NO: 136 to generate stably transformed plant cells, said polypeptide having tryptophan/tyrosine permease activity; and

b) selecting at least one plant from a population of plants derived from said transformed plant cells that has a difference in the level of oil in a tissue of said selected plant as compared to the corresponding level of oil in tissue of a control plant that does not comprise said nucleic acid.

54. The method of claim 53, wherein said difference is an increase in the level of oil.

55. The method of claim 53, wherein said polypeptide has the amino acid sequence set forth in SEQ ID NO:136.